

Bringing research back to life

Having all but killed it off, the British government is now brooding on the question of how best to restore a once-successful research enterprise to good health.

THE familiar cry of "What can be done to rescue British science?" seems about to be taken up by an unfamiliar voice — that of the British government, the chief source of the torments of the past decade. Even before next year's British budget had been made public (on Tuesday) by the new Chancellor of the Exchequer, Mr John Major, civil servants were brooding on what will happen a year from now.

The difficulty is that the answer to the question is no longer budgetary. If the value of research grants were forthwith doubled, that would not restore the research enterprise to good health. Nor would it help decisively if the funds available for research support allowed a greater proportion of those with bright ideas to succeed. What began as fiscal attrition (to be fair to the present government, in the early 1970s) became enforced decimation as research councils and universities were compelled to cut back on their establishments. Seen from the bench, the outstanding weakness of British science now is that there are too few good people competing for such resources as there are, and far too few outstanding research groups to which able ambitious people wish to attach themselves.

The seriousness of this state of affairs, however subjectively assessed, cannot be underestimated. The most productive and imaginative laboratories are those from which the most able researchers graduate in the greatest numbers. Rutherford's Cavendish Laboratory at Cambridge was a legend in its time, as was the Medical Research Council's Laboratory of Molecular Biology (still a first-rate laboratory) in the 1960s. Now, as then, the research web is constructed on a great diversity of centres — sometimes single people, sometimes an ingenious new instrument, sometimes the lucky coincidence of established senior people's different interests. Britain's trouble, now, is that it has too few centres of acknowledged excellence. Recent attempts to create them out of nothing (by means of the research councils' Interdisciplinary Research Centres, or IRCs) have been undermined by ambivalence about objectives.

The thing that British research most needs is a means of making sure that already outstanding research centres remain excellent. But which are they? The late University Grants Committee, now the Universities (no apostrophe) Funding Council, last year constructed a pecking order of university research departments that would have been even more revealing if overseas researchers had been

invigilators. Publicly supported laboratories have not been similarly assessed, but should be. The plan that this information should determine which centres compete in an improving way for the resources available will probably be frustrated by its complexity, and is certainly too slow. But there are other uses to which the data could be put.

The first need is to bless those centres identified as genuinely excellent with the resources that will allow them to remain internationally in the swim — travel money, stipends for visitors and the small change of competitive research of which there has been too little in Britain for too long. The simplest test of a centre's importance should be the number of PhD graduates from overseas who seek to work there (who should gratefully be helped to do so). Beyond that is the task of finding younger people whose talents promise excellence, and mechanisms by which they can also be supported generously.

But that can be only a means of stopping the rot. For the longer haul, there is a need for action on two different fronts. First, the profession of research must be made attractive again, which means fewer apparently arbitrary policy innovations (and reversals), less pressure to work in fields whose interest is externally defined and, uncomfortably for the John Majors of this world, much better pay. Second, there needs to be an understanding that excellent research is not merely worthwhile in itself, but necessary, in Britain now, if there are to be people skilled enough to prosecute applied research successfully. It is no longer a question of whether there is too much of one and too little of the other, but whether there is enough of either. □

Towards what unity?

German reunification makes necessary a broader integration on a wider front.

EAST Germany, as always, is a special case among members of the Soviet-led equivalent of the European Communities, called Comecon. For several decades, it has been the most productive member of its group. Now, after last Sunday's election, it seems certain to become, in a manner still undefined, one of the least advanced parts



of Western Europe. But not for long. The niceties of reunification apart, West Germany's government and business community seem determined that there shall be enough investment in East German industry to effect a rapid transformation of its industry — and to persuade East German families to stay where they are.

Meanwhile, this week's excitement notwithstanding, reunification is not yet a formality. One obviously outstanding question is that of security in Central Europe. Mr Mikhail Gorbachev's critics in Moscow, already uneasy that a buffer state should have elected to join the other side, are unlikely also to let the Soviet Union comply with Chancellor Helmut Kohl's assumption that a united Germany will simply be a larger member of the North Atlantic Treaty Organization (NATO). Yet the *status quo*, with hostile armies and missiles facing each other in the two parts of Germany, is plainly a nonsense. Willy-nilly, there will have to be a bargain, probably one in which nuclear missiles and foreign armies are removed. Is the West, or even Germany, ready for that?

The other complication is West Germany's membership of the European Communities. Even if the federal constitution allows eastern German-speaking states simply to accede, East Germany cannot quickly comply with community directives on matters such as water quality, highway safety and public subsidy of industrial production. But derogation of the rules in favour of what is now East Germany will set up a clamour by most other member states for exemption from what they regard as onerous obligations. It is not for nothing that M. Jacques Calvet, president of the French car-maker Peugeot, said at the weekend that the plan to create a true free market by 1992 had best be postponed until the uncertainties of German reunification have been settled. But that could mean postponement for good.

What can be done to overcome these difficulties? It would be easier if there were a coherent view of what the future of Europe might be like, but only Mr Vaclav Havel, now the president of Czechoslovakia, will chance his arm on that. He would dissolve NATO and the Warsaw Pact, and extend the European Communities to the East. A better bet would be that the two alliances should be merged, perhaps along the lines of the treaty that established Western European Union in 1948 (the crucial amendments to which were signed in 1954) and which, among other things, limits the freedom of acceding states to make nuclear weapons and long-range missiles. The same treaty requires its members to encourage the "progressive integration" of Europe, best accomplished in present circumstances by cultural means. Universities and grant-making research organizations have a larger role to play than they may think.

Outside Germany, East or West, politicians have been quick to wring their hands over the prospect of reunification, but have so far said very little about the ways in which Central Europe might accommodate the past year's dramatic changes in the East. What else, but reunification, can they expect? □

Market museums

The earliest fossil lizard has caught complacent British museums unawares.

MR Stan Wood, the freelance Scottish fossil-hunter, may have done more for British palaeontology than he knows — and more than British museums are yet willing to acknowledge. A few months ago, Mr Wood announced his discovery of the earliest known reptile, fondly known as 'Lizzie', and arranged to sell it to the *Museum für Naturkunde* at Stuttgart. Now, the National Museums of Scotland are scratching around to meet Stuttgart's price of £207,000, when British heritage legislation would give it the right to stand in Stuttgart's place. Inevitably, there are those who grumble that more patriotic fossil-hunters would not have put a British museum in that ignominious position. The truth is otherwise.

First, Wood has a clear legal title to his fossil and a right to sell it to whoever he chooses. Second, British museums should reflect that one of them would have had Wood's specimen for nothing if only they had persuaded him to work as a professional curator (which he once did). But, scandalously, there are no jobs. Last year, the Biology Curators' Group complained that 34 research posts have been axed from British museums since 1984, and that many more are vacant. And that is not surprising. A recent advertisement from the Liverpool Museum offered a top salary of £9,646 for a qualified biological curator. It is no wonder that only 33 natural history museums in Britain have qualified technicians to maintain their collections.

Why, it may be asked, when British museums cannot properly look after what they have, should they seek funds to match Stuttgart's bid? Or, to turn the argument around, if they had access to £207,000 without strings, would the purchase of the earliest fossilized reptile be the best use of the money? Chauvinism apart, complacent museum policies have conspired with a decade of public parsimony towards research to rob many institutions of a proper regard for their obligations.

Museums are also, of course, important means by which a wider public can share the fruits of professional scholarship, ultimately to the benefit of the professions concerned. The danger, especially at the British institutions forced by circumstances to levy admission charges, is that marketing takes charge. In Britain, the Natural History Museum (in London) is controversially in the van of those who regard their public image as a safeguard; its corporate plan, now imminent, will be scrutinized for signs that adequate salaries and staffing levels are being given the priority they deserve. Even in the marketplace museum, curatorship is not just a monetary sink. The museum that had employed Stan Wood would now have the earliest fossil reptile in a showcase at no cost — and would probably have earned £207,000 extra in admission fees. □

Many hands make light work of climate study

- 'Massively parallel' computers used
- Improved accuracy achieved

Washington

US researchers studying global change, frustrated by a shortage of computer time with which to run their huge and complex climate models, are turning to experimental, 'massively parallel', supercomputers as an alternative to expensive and heavily used traditional supercomputers such as those made by Cray Research.

A draft plan completed last week by the US Department of Energy (DOE) outlines a 10-year, \$240-million plus programme to adapt climate change models to massively parallel computers. Known as the Computer Hardware, Advanced Mathematics and Model Physics (CHAMMP) initiative, the project is subtitled "a DOE program for climate change". DOE has asked Congress for \$8 million in 1991 for the initiative.

Three US national laboratories have already started trying to convert existing climate change models to run on massively parallel machines. The parallel architectures of the new supercomputers allow them to carry out many simple tasks simultaneously, a feature that may apply well to climate models where the global climate is divided into hundreds or thousands of continuously changing cells.

Los Alamos and Oak Ridge National Laboratories are working on adapting atmospheric models, while the Lawrence Livermore National Laboratory will initially concentrate on the more complicated ocean models.

CHAMMP's eventual goal is to improve performance in climate change models by a factor of 10,000 relative to a single-processor Cray II. According to the draft CHAMMP report, a factor of 100 can arise directly from the switch to parallelism, a factor of ten can be achieved by fine-tuning the models' algorithms and another factor of ten should arise as compilers for parallel machines improve.

Massively parallel machines (sometimes known as 'hypercubes') may even be able to run typical climate change models a thousand times faster than an 8-processor Cray Y-MP — now the state of the art — within the next ten years, DOE researchers say.

By the end of the year, the DOE hopes to have equalled a Cray X-MP in running a standard atmospheric model designed by the National Center for Atmospheric Research (NCAR). The NCAR model is designed to run without modification on Cray 'moderately parallel' supercomputers, which have up to 64 processors.

Because the NCAR model is already designed with some parallel operation in mind, converting it to run on machines with hundreds or thousands of processors should provide a relatively straightforward test of massive parallelism in climate modelling.

Robert Chervin, who heads the NCAR team, describes the processor competition as "few-but-powerful versus many-but-weak". In contrast to the hand-wired Crays, which have a small number of extremely powerful and complex processors, the massively parallel machines string together hundreds of cheap off-the-shelf processors such as those found in desktop personal computers.

Initially, says Lawrence Livermore physicist Cecil Leith, "the main thing is to be cost-effective". Although today's massively parallel machines may not be as fast as a Cray at running climate models, they are many times cheaper, he says. One of the constant complaints of climate change researchers is the difficulty of finding enough computer time to run their models. Cheaper supercomputers should bring more power to all.

In congressional testimony last week, University of Illinois atmospheric scientist Michael Schlesinger told the House of Representatives sub-committee on science, research and technology that only one atmospheric research facility in the world — the West German Max Planck laboratory — has a dedicated supercomputer for climate modelling, although the UK Meteorological Office will soon be acquiring one as well. "Our hands have been tied. I can't think of another instance where the imbalance [between supply and demand] is as severe", he complained.

Present climate models divide the world into cells that measure 500 km on a side. Many scientists believe that it will be necessary to move to a higher resolution before basic questions about global warming can be answered. But a climate model with 50-km cells would take 40 years to run, Schlesinger said. Moving the models to massively parallel machines could be the answer. "The bang for the buck will be enormous if we can make it work", he said.

The largest hurdle so far in running climate models on massively parallel supercomputers is what Oak Ridge researcher Robert Ward calls "the load balance problem". "In one region you may be modeling a storm while in another part of the country you're modeling a

COLD FUSION

The party continues . . .

Washington

DESPITE the urging of a recent Department of Energy (DOE) panel against "any special funding" of cold fusion research, the department plans to double its budget next year for work in this field.

DOE will spend slightly more than \$1 million this year on five researchers who are studying cold fusion phenomena, according to Ryszard Gajewski, director of DOE's advanced energy projects programme. But because of continuing interest in the community, Gajewski says that amount is expected to increase to "about \$2 million" in the 1991 fiscal year, which starts in October; "there is high-quality basic research to be done".

"What we would like to see are carefully designed controlled experiments that would answer the question of whether these phenomena are really there. I have a completely open mind about which way it will be resolved." Late last year, an advisory panel convened by the DOE Energy Research Advisory Board concluded that experimental results had not presented "convincing evidence" that useful sources of energy would result from cold fusion phenomena. But although the panel recommended against the establishment of special programmes or centres, it was "sympathetic towards modest support" for cooperative experiments within the existing DOE funding system.

"We are following the recommendations of the panel", says Gajewski. New awards will go to cooperative research between "those who have seen effects and those who have not seen effects", he says.

Other current cold fusion funding sources include the state of Utah, which last year set aside \$5 million for the creation of a cold fusion centre around the work of Stanley Pons and Martin Fleischmann. And in June 1989, the Office of Naval Research also awarded Pons a \$400,839, two-and-a-half-year grant, of which more than \$240,000 remains to be spent.

G. Christopher Anderson

sunny day. You have to worry about keeping all the processors busy", he says. There may be a limit to how simple and numerous the individual processors in a parallel machine for climate models can be, he says.

NCAR's Chervin also warns that faster computers may not guarantee accurate models. "Some of the [scientific] issues cannot be beaten to death with more [computer] cycles. There's a major difference between scientific computing and computer science." The effect of clouds, water vapour and the oceanic absorption of carbon dioxide are issues that still resist accurate theory, no matter how much computer power is applied, he says.

G. Christopher Anderson

Kew Gardens get new exhibition hall

London

ARGUABLY the most energy-efficient public building in Europe opened on Tuesday (20 March) at the Royal Botanic Gardens in Kew, London. The Sir Joseph Banks Centre for Economic Botany cost £3.4 million to build and comprises a large exhibition hall, a library and enough space to store and study Kew's important reference collection, including some 9,000 glass jars containing wood samples. The centre will use no more energy than a family house.

The building works because most of it is underground. Cool in summer and warm in winter, it is ideal for the steady temperatures needed to maintain delicate specimens. Low-grade heat comes from a heat pump driven by two boreholes sunk 7 metres into the water table. In the winter, the ground water at this depth is warmer than air at the surface: the building design exploits this difference to provide warmth.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Interior view of the glazed concourse.

The radical design, one of 270 tendered, was the logical solution to a set of stringent design specifications. Apart from energy efficiency, the building had to blend in with several architecturally important buildings nearby. The exposed parts of the building are faced with fossiliferous Portland stone and crowned with a glass roof, similar in concept to the traditional Kew glasshouses as well as the modern, angular lines of the Princess of Wales Conservatory.

Economic botany has always been an important part of Kew's work, and the new centre is a tribute to the work of Sir Joseph Banks (1743–1820), the wealthy botanist who pioneered (and funded) expeditions to collect and study useful plants. Banks was the unofficial director of the Royal Gardens at Kew in the period before they opened to the public in 1840. He accompanied Captain James Cook on his voyage round the world (1768–1771), and organized the expedition to Tahiti to collect breadfruit trees for planting in British colonies. This latter expedition is remembered nowadays more for the name of its leader and his ship — *Bligh* and the *Bounty* — than for its botanical purpose. Henry Gee

Cross-licensing ordered

Washington

IN a decision that may put an end to the legal sparring between Amgen, Inc. and Genetics Institute over conflicting patent rights to erythropoietin (EPO), a judge in a US district court in Boston last week ordered the two companies to license to each other their respective versions of EPO, a kidney cell glycoprotein that stimulates red-blood-cell production. The agreements would be royalty-free.

Genetics Institute and its licensee, Chugai Pharmaceuticals of Japan, have been prevented from selling their drug, Marogen, in the United States because Amgen has been granted 'orphan drug' status for its version of EPO, Epogen. The court order asks Amgen to waive that status. Once Genetics Institute secures Food and Drug Administration (FDA) approval for Marogen, a process that should move quickly, both products can be sold in the US market. The move will allow the two companies to compete while leaving the door open for them to continue the patent dispute in the federal appeals court.

The judge said that implementation of the order could be delayed for 30 days to allow the details to be finalized and to permit "expedited appeals". Failure to comply with the court's order would result in a permanent injunction enjoining the company concerned from infringing the other company's EPO patent.

PHARMACEUTICAL INDUSTRY

Amgen has also been asked to get in touch with the appropriate administrative agencies, which Genetics Institute takes to mean the US Food and Drug Administration (FDA) and the International Trade Commission, and to withdraw all opposition to the efforts of Chugai to introduce Marogen to the US market.

Although the FDA has publicly stated its intention to approve Marogen, its rival's orphan-drug status has proved a greater stumbling block to Marogen than expected. The orphan-drug law provides manufacturers with incentives, including a seven-year marketing monopoly, to develop new treatments for rare diseases.

But Genetics Institute believes that the market for EPO is so lucrative that there is ample incentive for more than one company to compete in the marketplace.

Although the order may provide an agreement in the short term, it does not prevent one or both parties appealing against the original patent ruling of 11 December 1989 in the federal appeals court, or from claiming damages for patent infringement. Last December, a Boston court upheld the central claims of both companies' EPO patents (Amgen's US patent: 4,703,008; Genetics Institute's US patent: 4,667,195), while also declaring them partially invalid and mutually infringing (see *Nature* 343, 500; 8 February 1990).

Diane Gershon

Another drug-house merger

Washington

THE year-long global consolidation of the pharmaceutical industry continued last week with the announcement that Rhone-Poulenc SA, France's largest pharmaceutical company, is to merge its human prescription-drug business with that of the Rorer Group, Inc. of Philadelphia. The new company will be called Rhone-Poulenc Rorer, Inc., 68 per cent of which will be owned by Rhone-Poulenc. Unusually, Robert E. Cawthorn, chairman and chief executive officer of Rorer, will retain those responsibilities in the new company.

In the deal, Rhone-Poulenc will purchase just over half (21.6 million shares) of Rorer's outstanding common stock at \$78 per share. It will also issue a number of contingent value rights (CVRs) equal to the number of remaining outstanding common shares held by Rorer's public shareholders. The CVRs will entitle shareholders to a payment from Rhone-Poulenc if the market price does not reach a target value of \$98.26 within three years of the merger.

The merger will cover only Rhone-

Poulenc's human pharmaceutical business, and will not affect serum and vaccine production, veterinary medicine, animal nutrition and Rhone-Poulenc's interest in Roussel Uclaf SA.

This announcement follows a series of mergers in 1989 that included the formation of such pharmaceutical giants as SmithKline Beecham and Bristol-Myers Squibb (*Nature* 338, 610; 1989). Last week, SmithKline Beecham announced it is to cut its worldwide workforce of 55,000 by ten per cent, sparing only its sales force and drug research and development.

SmithKline Beecham's pre-tax profits of \$1,168 million, up three per cent from 1988 when the companies existed separately, were coolly received on Wall Street. Many analysts feel that interest payments on SmithKline's \$3,100 million debt are responsible for the flat profits. Countering these claims, Robert P. Bauman, SmithKline Beecham's chief executive officer, insists that the company is right on track, but acknowledges that a major objective "must be to reduce our debt and interest costs".

Diane Gershon

DEEP DRILLING

Japan plans holes on land . . .

Kyoto

JAPANESE geologists are planning to drill a hole 15 km deep into the Earth to penetrate the boundary of a plate that is subducting beneath Japan.

The Tartarus project, named after the Greek name for the abyss below Hades, began last month with a survey of potential sites by the Geological Survey of Japan and the Disaster Prevention Centre. The project is expected to take ten years and \$500 million to complete.

The best candidate for the deep drilling site is the seismically active Izu peninsula, south of Tokyo, where the Pacific and Philippine plates plunge beneath the

Eurasian plate on which Japan stands. One aim of the project is to study the region of the Earth's crust where deformation changes from brittle to ductile. In the Izu peninsula, this region lies only a few kilometres below the surface and some of the most destructive earthquakes in Japan develop there.

The survey involves both seismic and magnetic studies and has a budget this year of ¥30 million (\$200,000). Several different sites will be studied including the Izu peninsula. Work is now under way in Hokkaido and the Kanto region near Tokyo.

The Science and Technology Agency (STA) says that it is interested in providing funds for the deep-drilling project, although it cannot say when backing will be available. The Ministry of International Trade and Industry (MITI), to which the Geological Survey is affiliated, is also interested in deep drilling because it may help to provide a new energy source. Scientists at the University of Tokyo have recently suggested that methane in natural gas fields in Japan may be directly derived from the Earth's mantle (see *Nature* 305,

792; 1989).

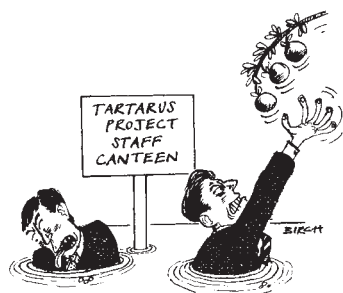
Japan, and the Izu peninsula in particular, pose considerable technical difficulties for deep drilling because the region is so seismically and volcanically active. The deepest hole in the world is in the Kola Peninsula, where Soviet scientists have drilled to a depth of 13 km. But German plans to reach this depth in the Oberpfalz were abandoned after a pilot hole of 4 km showed unexpectedly high temperatures.

Simon Wallis

. . . and in the ocean

Japan's Science and Technology Agency (STA) hopes to build an ocean-going vessel for deep-sea drilling, similar to the US JOIDES ship and capable of drilling through the deep-ocean floor to a depth of 4 km. The STA ship would be quite independent of the international Ocean Drilling Program (ODP), which is at present partly funded by the Ministry of Education Science and Culture (MESC). The proposal to build a new ocean-going drilling vessel was made by the advisory committee to the STA ocean development section and a feasibility study with a budget of ¥13 million (£87,000) will begin in April.

Simon Wallis



COMPUTER NETWORKS

US launches transatlantic links

Washington

BUOYED by an exponential rise in traffic on its US computer network, the National Science Foundation (NSF) last week launched the first high-speed scientific computing link to cross the Atlantic. With the new line, NSFnet now has a full-speed link to the European Academic Supercomputing Initiative network (EASInet). Operating at 1.5 megabits per second (known as T1 speed), the fibre-optic link is about 24 times faster than previous inter-continental connections.

The European Laboratory for Particle Physics (CERN) in Geneva serves as the European 'gateway' for the link, which is connected at its other side to the NSF supercomputer centre at Cornell University in Ithaca, New York.

NSF's networking partners (IBM, MCI Communications and the Merit university consortium) also demonstrated the first prototype 'T3' link at a trade show in Washington last week. At 45 megabits per second, the T3 link is 28 times faster than the T1 connections.

NSF intends to begin upgrading the NSFnet 'backbone' — nodes at seven US universities and all six NSF supercomputer centres — to T3 speed later this year.

Figures released by NSF this month show that traffic on NSFnet has increased 445 per cent in the last year, to 2,500 million 'packets' a month. (A packet is a variable amount of data with some support-

ing information appended, such as a destination and routing symbols. Although packets can vary from 48 to 4,000 bytes, NSFnet packets statistically average 125 bytes.) The number of NSFnet users is expected to grow from about one million today to between four and six million — approximately one half of the US academic and research community — by the end of the decade. In congressional testimony last year, NSF officials warned that without new networking initiatives the rapidly climbing volume of network traffic would soon swamp NSFnet.

But new users alone are unlikely to inundate a T3 — or even a T1 — network in the near future, says Michael Conners, director of computing systems for IBM Research. Instead, traffic is expected to grow as scientists react to the increased responsiveness and reliability of the new high-speed links by expanding their network use from electronic mail to interactive operations of remote computers and the exchange of large amounts of graphic data.

Although electronic mail is still the largest single use of NSFnet, interactive applications (now 20 per cent) and file exchange (29 per cent) are gaining ground. Further network capacity, in the form of more T3 links, is sure to spawn "a whole new set of applications", including the interactive 'steering' of scientific simulations on remote supercomputers, Conners predicts.

G. Christopher Anderson

MARINE POLLUTION

UN report says concern is wrongly directed

London

PUBLIC and government concern over marine pollution should be refocused, according to a report on the state of the marine environment, sponsored by the United Nations and published this week. The report, by 20 marine scientists led by Professor Alisdair McIntyre of the University of Aberdeen, concludes that oil spills and heavy metal contamination pose less of a threat than was thought in the early 1980s.

Launching the report in London, McIntyre said that over-enrichment — eutrophication — of coastal waters from nitrates and phosphates in run-off from agricultural land is now a major concern, and pointed to recent algal blooms in the Dutch Wadden Sea. In tropical coastal waters, there is the additional problem of a build-up of chlorinated hydrocarbon pesticides, still used extensively in developing countries, he said. But despite problems near coasts, McIntyre said the open oceans remain largely unpolluted habitats.

He also suggested that much-publicized moves to reduce sewage dumping at sea may be misdirected. McIntyre believes there may be scope for increased disposal of treated sewage in deep waters.

Peter Aldhous

US joins Soviet venture

Washington

A convergence of chance events, including the death of physicist Andrei Sakharov and a Soviet decision to adopt international radio frequencies for its space projects, has led to a new US-Soviet co-operative space venture.

Earlier this month, US Vice President Dan Quayle announced that the United States would provide ground-based hardware and services for a Soviet radiotelescope to be launched in the mid-1990s. In a joint appearance in Washington, Quayle and Yri Dubinin, the Soviet ambassador, applauded the venture as a "significant step" towards new cooperation.

Known as RADIOASTRON, the 10-metre orbiting Soviet telescope will work with ground-based radiotelescopes to form a very-long-baseline interferometry (VLBI) network of unprecedented size. Present VLBI networks are restricted by the Earth's diameter to several thousand miles, but with one element in Earth orbit, the baseline can be expanded by an order of magnitude. The RADIOASTRON network, like other VLBI networks, will be used to make high-resolution maps of distant galaxies, quasars and pulsars.

One of the last letters signed by Sakharov before his death was an appeal to the United States to collaborate in the Soviet venture. Although officials at the National Aeronautics and Space Administration (NASA) and Soviet officials have discussing joint participation in

RADIOASTRON for three years, talks had started on equipment compatibility and export controls, says Samuel Keller, NASA assistant deputy administrator.

But in recent months the US National Radio Astronomy Observatory has at last received permission to export four special data recorders built for the Soviet mission, RADIOASTRON planners agreed to switch to communications frequencies that are compatible with US and international facilities, and Sakharov's stepson returned from the Nobel laureate's funeral to deliver the posthumous appeal to NASA administrator Richard Truly. The convergence of events adds "a special note of poignancy" to the cooperative venture, Quayle said.

In addition to the data recorders, which are on extended loan, NASA will provide data-processing services and will track RADIOASTRON with its Deep Space

Network, which is an order of magnitude more accurate than Soviet networks, according to Keller. The additional worldwide tracking will allow RADIOASTRON to make useful observations 90 per cent of the time, more than twice what it can do with Soviet tracking alone. The total cost to the United States is expected to be \$20 million over five or six years.

The Soviets are building a ground-based 70-metre radiotelescope in Uzbekistan to support the RADIOASTRON project. Known as the Suffa telescope, the new facility will join two existing Soviet 70-metre radiotelescopes in the Crimea and the Far East.

A group of 70 scientists from 15 countries is additionally planning a 25-metre space radiotelescope known as the International VLBI Satellite (IVS). The IVS is "a natural next step" to the RADIOASTRON mission, wrote Sakharov, adding that "we hope that NASA will support this mission also".

G. Christopher Anderson

SPACE SCIENCE

Financial hiatus for Hipparcos satellite

London

THE European Space Agency (ESA) council this week debates future budgetary increases for the Horizon 2000 space science programme, following the presentation of a report on Horizon 2000's management and troubled finances. The programme, set up in 1984, has been plagued by Britain's reluctance to increase contributions in line with its partners — a major problem, as Horizon 2000's budget must be agreed unanimously by the 13 subscriber states. After a year of negotiation, Britain agreed the last Horizon 2000 budget increase in December 1988 only when ESA commissioned a review of the programme.

The review team, chaired by Professor Klaus Pinkau of the Max Planck Institute for Plasma Physics, near Munich, has not yet made its findings public, and subscriber states' delegations are unwilling to reveal their negotiating positions on Horizon 2000's budget. But the British delegation is unlikely to agree funding for specific space science projects until problems raised in the Pinkau report have been debated in full.

The Hipparcos astronomy satellite is one project caught up in ESA's continuing budgetary arguments. Following its failure to reach a geostationary orbit after launch (see *Nature* 340, 491; 1989), Hipparcos now communicates with three ground stations, rather than one. This has increased costs so that Hipparcos's original budget will run out in July.

Project scientist Michael Perryman says that a "silent success" has followed the mission's well publicized early difficulties, and Hipparcos is expected to have a further two or three years' life. The satellite will

then collect most of the star-mapping data originally hoped for, but this will need an extra 26 million AU (accounting units; 1 AU = £0.67).

ESA's Science Programme Committee discussed Hipparcos's funding last week. Peter Smith, from the British National Space Centre (BNSC), says opinion was "fairly evenly divided" on whether to agree funding for Hipparcos in advance of the Pinkau report. But the British view prevailed, and a final agreement may now have to wait until June.

Smith says that Hipparcos will almost certainly get the necessary funding. The uncertainty is where the money will come from. The Horizon 2000 budget cannot easily accommodate unforeseen expenditure without cutting back, so ESA may wish to ask subscriber states for more money. Britain "would find it extremely difficult" to produce an extra Hipparcos contribution, Smith says, and might not be alone in this.

Britain's problems with space science funding stem, in part, from the absence of a separately funded national space agency. BNSC has only a coordinating role, and space science is funded from the Science and Engineering Research Council's budget.

Contributions to ESA's science programmes must compete with unrelated research areas, in addition to Britain's space science efforts outside of ESA. British space scientists are keen to collaborate with the US National Aeronautics and Space Administration on ultraviolet astronomy projects, where ESA has no future plans.

Peter Aldhous

ANIMAL RESEARCH

Monkey imports may be curtailed in US

Washington

IN the wake of the discovery last year that monkeys in two US laboratories are carrying the deadly Ebola virus, health officials are considering banning imports of the main species of research primates.

In a letter to about 200 primate importers, William Roper, director of the US Centers for Disease Control (CDC) warned that "active transmission [of Ebola-like virus] in monkeys in quarantine has been demonstrated. The health of your workers may be at risk." Antibodies to the virus have recently shown up in the blood of two US laboratory workers.

CDC are considering temporary restrictions on the importation of cynomolgus monkeys, which make up 75 per cent of all primates imported to the United States. Although about 16,000 cynomolgus monkeys were imported last year, many large primate centres also breed their own and should be able to weather temporary import restrictions.

G. Christopher Anderson

Plague of satellite failures

Washington

Two US space mishaps have left a new \$1,000-million spy satellite in fragments and a \$150-million communications satellite in a tenuous orbit waiting for a rescue that may never come.

On 14 March, a seemingly perfect launch of an Intelsat telecommunications satellite ended on the verge of disaster when faulty wiring in a Titan III booster prevented the satellite from separating from the second stage of the rocket. Although Intelsat was able to free the satellite by jettisoning an unfired 'kick' motor attached to the defective second stage, the bus-sized vehicle was left stranded in a low elliptical path without the means to reach its intended orbit 22,300 miles above the Earth.

With the help of the Defense Department's radar tracking network, Intelsat was eventually able to re-establish communications with the satellite several hours after the launch.

By the end of the week, project managers had successfully fired thrusters on the satellite that boosted it from an orbit that would have brought it back to Earth within two weeks to a slightly higher orbit that will give the communications company several months in which to plan a rescue.

But it is not clear if such a rescue will be possible. A National Aeronautics and Space Administration (NASA) spokesman says that it is "premature" to talk about conscripting the Space Shuttle to recover the uninsured satellite. Shuttle missions typically cost over \$300 million, although a rescue appended to the end of another, previously scheduled, mission would be substantially cheaper. One possible opportunity to add a rescue effort to a shuttle mission is during the scheduled launch of a gamma-ray observatory in November, according to John Pike, a Federation of American Scientists space analyst.

An independent company, Orbital Sciences Corporation, has written to NASA and Intelsat offering to carry one of its own rockets aboard the shuttle, strap it onto the stricken satellite and boost the vehicle to its intended orbit without need to return it to Earth for relaunching. NASA had no immediate comment on the proposal.

Repercussions from the failure are likely to be felt most in the infant US commercial space business. Martin Marietta Corporation, the aerospace company that built the Titan III rocket, is the weakest of the three US companies (General Dynamics and McDonnell Douglas are the others) that provide commercial launch vehicles.

Since 1985, the company has suffered

five failures of varying severity in nine Titan III launches, although a company spokesman points out that the rocket has a "90 per cent success rate" in its 25-year history. A one-in-ten failure rate is taken for granted in the space business. But Martin Marietta's recent failure rate of more than 50 per cent is either "fiendishly bad luck" or indicative of widespread production problems, says Pike.

The latest mishap may accelerate Martin Marietta's departure from the commercial space business. The company has already converted its main assembly line to build new Titan IV boosters for the Air Force and NASA (see *Nature* 339, 567; 1989) and closed its Commercial Titan office in Washington. The only remaining scheduled Titan III launches are another Intelsat satellite in June, and the Mars Observer mission for NASA in September 1992. Titan IIIs are "too big and too expensive" — at \$115 million each — for commercial launching, says George Washington University analyst John Logsdon, who thinks the latest malfunction may be "the nail in the coffin" for the company's commercial space venture.

The Intelsat mishap is the second commercial launch failure in three weeks. Ariespace, the European consortium,

lost two Japanese telecommunications satellites at the end of March when its new Ariane-4 rocket exploded (see *Nature* 344, 7; 1 March 1990).

Late last week, the US Defense Department revealed that the \$1,000 million spy satellite that had been placed in orbit by the space shuttle two weeks ago had broken up in space. Few details were available to explain why or how the vehicle malfunctioned. The *Washington Post* reported that Soviet space officials were tracking four large sections of the disintegrating satellite and were predicting that the fragments would re-enter the atmosphere over northern Soviet territory before 15 April.

The Department of Defense said: "Space shuttle mission STS-36 achieved its goal associated with a classified [Defense Department] mission. Hardware elements associated with the mission are expected to re-enter the Earth's atmosphere. We believe there is no risk to life or property." Initial speculation was that the satellite has been spinning out of control for some time because of problems with its control system. Centripetal force or an internal explosion could have caused it to disintegrate. A prototype version of the satellite was launched last August and tumbled for weeks before ground controllers were able to stabilize it.

G. Christopher Anderson

SCIENTIFIC SOCIETIES

German physical societies plan to reunite

Munich

THE German Physical Society (Deutsche Physikalische Gesellschaft or DPG), one of the oldest and most respected scientific societies in the world, announced last week that it will be reunited as soon as possible with the Physical Society of East Germany. The new DPG will have 20,000 members, about 90 per cent of them from the West.

DPG is the first scientific institution to be 'reunified' since the change of government in East Germany. Founded in 1845, DPG had an illustrious history until the end of the Second World War and the division of Germany, when the East German society split off.

DPG leaders say that the quick unification was only possible because parallel organizations continued to exist in the two German states. Other professional organizations, such as the Association of German Engineers (VDI in German), were banned by the East German government.

The announcement came at the annual meeting of the DPG in Munich, where 200 East German physicists attended as guests of the DPG. DPG began free mailings of the DPG proceedings, *Physikalische Blätter*, as soon as the Berlin Wall came down last fall.

The new DPG will begin officially to hold

joint meetings and other activities as soon as the legal details are worked out.

The newly elected president of the East German physical society, Gerd Röpke of Rostock, said that an overwhelming majority of members voted to join DPG at the first opportunity. The East German physical society is at present subordinate to the East German Academy of Sciences, which itself has an uncertain future.

DPG President Otto Folberth said an important task for DPG will be to try to prevent a further 'brain drain' from East Germany. In addition to providing copies of *Physikalische Blätter*, Folberth said DPG was trying to organize donations of books and research materials to East Germany.

Röpke was optimistic that East German physicists can be persuaded to stay in what is now East Germany as the material situation for research improves. Collaborations have already sprung up in Berlin and other places in areas such as semiconductor research.

West German physicists present at the meeting said they saw great potential for fruitful collaborations with East German researchers, especially in areas such as solid state physics, polymer research and atmospheric physics. Steven Dickman

Campaign to shut plant

Greifswald, East Germany

THE fate of East Germany's controversial Greifswald reactor complex is likely to be decided next month, when an investigatory commission reports to the East and West German governments. The commission was set up at the urging of the West German Environment Minister after the news magazine *Der Spiegel* revealed in January that one reactor had narrowly avoided a catastrophic accident in 1975.

Some experts at Greifswald believe the Soviet reactor design is fundamentally flawed. If they are judged correct, a shut-down could have dramatic repercussions in three other Eastern European countries where ten reactors of the same type are in operation. Four similar reactors are planned for Greifswald.

Two of the four operative 440-MW reactor blocks at Greifswald were temporarily shut down in February, on the advice of Western experts. The other two blocks continue to operate at lower power than normal, providing both power and heating for the nearby town of Greifswald (population 75,000) and environs.

Physicist Norbert Meyer, who heads a research group in materials science at the plant, says that the plant should be closed because of serious design flaws, many of which have already been identified in a preliminary report from the West German Society for Reactor Safety (GRS). Meyer says that he "cannot rule out that a serious accident might occur" at Block 1. Block 4 should be operated only as long as there is no other source of heat for the city of Greifswald, he says.

East German minister without portfolio Sebastian Pflugbeil of the New Forum says it is "irresponsible" to continue to operate the reactors. Pflugbeil is a member of the so-called 'round table' set up to oversee the transition government until this week's East German elections. As soon as the accident and continuing safety problems at the plant were revealed, he introduced a resolution at the round table to shut the plant. The government has not responded.

The most serious flaw is that the main and auxiliary control cables run in the same ducts and in some cases are not even redundant. In the 1975 accident, a fire destroyed the cables providing normal and emergency power. Only through a lucky accident — a single pump hooked up to an outside power supply — were plant operators able to avoid a meltdown and provide cooling water to the reactor for several hours until reserve power lines were installed.

The reactors are not protected against breaks in the pipes that deliver cooling water. A break in even one of the thin water lines, greater than 32 mm in dia-

meter, could result in a meltdown. Western reactors are protected against breakage of cooling lines up to the largest pipes, which can reach 800 mm in diameter.

After 25 years of bombardment by neutrons, the steel walls of the reactor pressure vessels are dangerously brittle. This was the main reason that Blocks 2 and 3 were shut down last month. A break in the vessel, which contains water pressurized to 125 atmospheres, would lead to a meltdown.

Block 1, which went on line in 1973, reached a level of brittleness far above the margin of safety by 1982. But it took until 1988 before Block 1 was finally shut down for 'thermal annealing' — heating to 480 °C for 120 hours to restore the tensile strength of the steel.

Block 2 passed the maximum safe level of brittleness in 1988, but was not shut down until earlier this year. Block 3, which went on line later — in 1977 — after modifications had been made in the design, was within a tiny fraction of the maximum when it was shut down in late February. Heinz-Peter Butz of GRS said it was "very hard to understand" why the plant operators did not shut down the blocks sooner.

The Soviets recognized the gravity of the problem in the mid-1980s and hurried to build a thermal annealing apparatus that could be used for all reactors of this type. But both Meyer and GRS say that it

ENERGY POLICY

Boston utility to foot bill

Boston

IN a \$215-million plan that is the first of its kind, Boston Edison, one of New England's largest utility companies, will at least temporarily forestall the need for new power plants by paying instead for an array of energy conservation measures. The conservation programme, negotiated by representatives from the utility, state government and local environmental groups, is hailed by all three groups as a "revolutionary shift" and "a model of enormous importance for the rest of the nation".

The plan is expected to reduce the demand for electricity in the Boston area by a thousand million kilowatt hours over the next five years, roughly equivalent to the amount of power generated by a small 120-MW power plant. It is particularly notable for its timing, announced shortly after the Seabrook nuclear power plant in neighbouring New Hampshire was cleared finally for its full-power licence after a gruelling and costly decade-long ordeal (see *Nature* 344, 96; 8 March 1990).

Boston Edison will invest immediately in measures such as making new buildings

has not been proved that the annealing process restores the necessary tensile strength.

Many thousands of the tiny 'needle tubes' that are part of the steam generator have been severely corroded, according to Meyer, making the reactor vulnerable to a 'pressure wave' somewhere in the secondary cooling circuit. Such a wave, which could have many causes unconnected to the reactor itself, would again carry the risk of a possible meltdown.

Spokesman Peter Beier of the Greifswald plant says that "we don't claim that these reactors meet the safety standards of the 1990s", but all the nonsense about "time bombs" and "Chernobyls" that has appeared in the Western press is "horrendous nonsense". "We see no alternative" to rebuilding at least two of the four reactor blocks at Greifswald, said Beier. Despite the threat to the plant, the Greifswald management has not yet put forward an alternative plan for providing heating and electricity to the town.

Monika Litwin, spokeswoman for the Heavy Industry Ministry responsible for Greifswald, said that the plant "never caused any immediate danger to human lives", either in 1975 or at any other time.

After Meyer made public his charges, he had to face 800 of his co-workers at Greifswald, who demanded his immediate dismissal. Although he has been kept on, Meyer is no longer allowed into sensitive areas at the plant. "I am not considered reliable any more", he said. **Steven Dickman**

more energy-efficient than present standards require, retrofitting existing residential and commercial buildings and supplying energy-efficient light bulbs free or at low cost.

During the next five years, the utility expects to save two dollars' worth of electricity for every dollar invested. In an innovative economic incentive system, the utility will be allowed to retain a portion of this expected \$400 million in energy savings as profit in addition to recouping its costs. The rest of the savings will be passed on to customers. With such a market incentive, conservation becomes equally if not more profitable for the utility as the construction of new generating capacity.

Although conservation plans are not new, observers say that the Boston Edison announcement marks a trend towards a distinctly new movement on the part of electric utilities towards conservation. The intense opposition to the Seabrook nuclear power plant may be an extreme case, but is indicative of the difficulty of bringing any new power plants on line.

Seth Shulman

CONSERVATION

Big new forest reserve

Cape Town

SOUTH Africa's Minister of Environment Affairs, Gert Kotze, announced last month the establishment of a huge new conservation area on the northernmost part of the country's east coast. It will bring together the existing Lake St Lucia and Mkuze game reserves, the Sordwana Bay National Park, several state forests and the St Lucia Marine Reserve, and will include a state forest on the eastern shores of Lake St Lucia where a controversial dune mining operation for titanium has been proposed. Last week in parliament, Kotze denied that the declaration of the new conservation area was a "trade-off" for granting permission for the development, and said that the government is awaiting the outcome of an environmental impact assessment before making a final decision on the mining proposal.

The Lake St Lucia Game Reserve is, together with the Umfolozi and Hluhluwe Reserves, the oldest in Africa, having been proclaimed in 1895. Lake St Lucia is the largest estuarine system on the continent, and contains a variety of wetland and other habitats considered critical to the survival of 102 rare and endangered species. But since 1954 it has been damaged by the introduction of large-scale timber plantations, which, together with water abstraction from its main source, the Mkuze River, have led to serious hypersalinity problems. These problems should soon be alleviated, as the new reserve will incorporate much of the Mkuze catchment zone, and areas under timber plantation will be replanted with indigenous vegetation.

In addition, a South African Defence Force missile-testing range at St Lucia is to be phased out. The consolidated area will be more than 275,000 hectares in size, making it the third largest reserve in South Africa. But the proposed dune-mining operation has become perhaps the country's most controversial development to date: a petition opposing it has drawn over 250,000 signatures. Prospecting and subsequent mining rights to the area were granted about 20 years ago to Richards Bay Minerals, 50 per cent of which is owned by the British-based multinational Rio Tinto Zinc company. The mining operation would bring in an estimated R5,000-R9,000 million (\$1,250-\$2,250 million) over its 20-year lifespan, and create 600 jobs in an area of high unemployment.

Richards Bay Minerals argues that 60 per cent of the 1,300 hectares due to be mined is covered by exotic pine plantations which will be replanted with natural vegetation once the operation is finished. Environmental pressure groups have countered that their main concern is less

that a particular area will be mined than that power lines will have to be built across the lake, and that trucks filled with ore will travel through the reserve at a rate of one every five minutes for 24 hours a day. The National Union of Metalworkers of South Africa has been reluctant to take sides, but has expressed reservations about the way that the campaign for St Lucia has been run. "None of the environmental organizations have consulted us about the issue and some of our members are wondering if they think it is more important to save insects and animals while people have to sacrifice jobs and wages", their regional secretary, Michael Mabuyakhulu, was quoted as saying in the *Weekly Mail*.

Michael Cherry

RADIATION EXPOSURE

More haste, less speed

London

IN its anxiety to show that follow-up studies to the recent Gardner report are in hand, British Nuclear Fuels Limited (BNFL) has trodden on the toes of one of Britain's leading epidemiologists.

Professor Peter Smith, from the London School of Hygiene and Tropical Medicine, says he was telephoned by Roger Berry, chief medical officer for BNFL, and asked to give an independent assessment of the Gardner report. This report suggested a link between the cases of childhood leukaemia near BNFL's Sellafield nuclear plant and their fathers' occupational radiation exposure (see *Nature* 343, 679, 22 February 1990).

Smith sent BNFL some preliminary notes before visiting a research project in Africa, but says he stressed these "were not for general distribution or release". Shortly afterwards, BNFL issued a press release which summarized Smith's preliminary report, describing him as "BNFL's epidemiology consultant".

Smith says BNFL's behaviour was "inappropriate", and is particularly distressed by its description of his status. Although he has worked on BNFL-commissioned research projects in the past, Smith says he is an independent scientist and was not paid to advise BNFL on the Gardner report. A spokesman for BNFL said that its action was in line with a commitment to keep the Sellafield workforce fully informed (the press release was based on a letter sent to trade union representatives). But he added "we are sorry if he [Smith] is unhappy" and conceded "it may well have been better" to have described Smith's status differently.

Smith has now issued his own press

AIDS

Mann's resignation causes uncertainty

London

THE World Health Organization (WHO)'s Global Programme on AIDS faces an uncertain future following the resignation of its founding director, Jonathan Mann, last week. In his resignation letter to WHO director-general Hiroshi Nakajima, Mann cites the "great variance" between his own and Nakajima's positions on important areas of policy. Nakajima and Mann are understood to have disagreed, among other issues, on the programme's campaign to highlight the problems of AIDS in women, and links with other non-governmental organizations.

Mann's resignation has saddened AIDS researchers. Geoffrey Schild, of the UK Medical Research Council's directed programme on AIDS, describes Mann as "a very vigorous and effective campaigner on public health policy".

The choice of Mann's successor will have an important bearing on the future direction for the programme, which spends more than \$100 million a year, coordinating research and international efforts to slow the spread of the disease. Schild says "the leadership of WHO in this field is imperative".

Most recently, the programme has become involved with the development of AIDS policy in Eastern Europe (see *Nature* 344, 95; 1990). Mann will remain as director until June.

Peter Aldhous

release which summarizes his views of the Gardner report. He concludes that the study was "well planned, well executed and well analysed". Some problems with the quantity of obstetric and questionnaire data were not surprising, given the 36-year timespan covered by the study. Smith writes "it would be prudent to assume the association with external radiation is causal" until more evidence is found, although he cannot rule out a chance association, or the effects of exposure to chemicals. He adds that further studies are needed on other groups of radiation workers.

BNFL also asked Sir Richard Doll, of the University of Oxford, to review the Gardner findings, and he has now reported back.

Peter Aldhous

■ LAST week's article on US reaction to the Sellafield childhood leukaemia study mentioned that the American Statistical Association (ASA) will have a session on radiation epidemiology at its August meeting in California; ASA has also scheduled a special conference on radiation and health in Copper Mountain, Colorado on 8-12 July at which Professor Martin Gardner, who headed the Sellafield team, will speak.

G. Christopher Anderson

Research with fetal tissue

SIR—Keith Crutcher's bizarre letter on research with fetal tissue (*Nature* 343, 10; 1990) has prompted me to join this debate. Abortion has a negative moral import to most people, and most would also agree that the gravity of offence rises to that of murder by the end of gestation. Before that, however, a pregnant individual may weigh the moral costs of abortion against the moral costs of having an unwanted child. These can include the effects on the mother's life, the effects of abuse and/or inadequate support on the child and the legal and economic costs visited upon society by the adult that emerges. Because such moral calculations vary greatly among reasonable people and societies, decisions on abortion are best left to the individuals involved. Laws are one means by which society promulgates common moral beliefs. The historical and current lack of a moral consensus on abortion is exactly why abortion is not now illegal.

The same logic may be applied to the question of research with fetal tissue. As Crutcher admits, the legality of abortion in effect legalizes the existence of fetal tissues and their availability for research. What happens to them after that should be decided by the individual scientist. Those with moral objections either to abortion or to the use of fetal tissue need not accept such work personally; others may decide that the moral benefits obtained from such research outweigh the moral costs. The scientific community generally favours research with fetal tissue and human embryos; the specific ban on its funding by the US Department of Health and Human Services is therefore inconsistent with society's lack of consensus on abortion or fetal research, as well as with the opinions of most individuals in this community. As we all know, funding decisions on research projects are best made on the basis of peer review of their scientific merits.

The possibility of finding a moral consensus on abortion in the future is only decreased by the polarizing language and tactics adopted by those at either extreme, including Crutcher's deliberate attempt to provoke sensory disgust. It stands to reason that this lack of consensus necessarily favours the pro-choice side. Anti-abortionists should therefore be more motivated to compromise, and should abandon untenable, absolute positions. One possible compromise is for elective abortions to be legal only in an earlier and shorter period of pregnancy, such as 0–13 weeks. This preserves most of the rights of a mother to choose, but clearly segregates the early embryo or fetus from those near separate viability (25 weeks). In addition, anti-abortion groups may want to devote

their energies to preventing unwanted pregnancies in the first place.

ROBERT BOOKSTEIN

Department of Pathology M-012,
School of Medicine,
University of California, San Diego,
La Jolla, California 92093, USA

Shockley defamed

SIR—That *Nature* saw fit to publish such abusive treatment of the late William Shockley, not once but twice, strikes me as a disgrace.

First, the obituary on Shockley (*Nature* 341, 190; 1989) is clownishly signed with the obvious pseudonym "H. Kallikak" (harking back to the ancient "Jukes" and "Kallikaks" studies of 'feeble-mindedness' by Dugdale and Goddard). Scarcely 12 per cent of its wordage is devoted to Shockley's path-breaking achievements in physics (for which he won the Nobel prize), while all the rest is a grossly opinionated disparagement of what the authors refers to as the "affliction" Shockley developed after becoming a laureate, namely his interest in IQ, genetics, population differences and their possible long-term social consequences. As one who has paid close attention to everything Shockley published on these topics for more than 20 years, I must say that in my opinion 'Kallikak's' comments give a muddled and biased view of Shockley's thinking in this area.

Second, Frederick Seitz's pretended defence of Shockley (342, 474; 1989) only adds injury to insult, essentially going along with 'H. Kallikak' but excusing Shockley's "ill-conceived concentration on socio-genetic matters" as possibly the result of an automobile collision, a conjecture that leads *Science's* capsule summary (247, 25; 1990) of Seitz's letter in *Nature* to question: "Could William Shockley's unpopular ideas about race and IQ have been the result of a blow to the head?" This surely must be the farthest outpost of *ad hominem* criticism of a scientist's ideas. As one who knew Shockley well during the past two decades of his life, I can testify that, in matters analytical, mathematical, statistical and scientific, Shockley was a remarkably clear thinker and absolutely as sharp as they come. I question that anyone who had any intellectual commerce with Shockley during this period could have any doubt that he was immensely bright and quick of mind.

'Kallikak' and Seitz both seem unaware that their sweeping dismissal of the basis of Shockley's ideas on IQ, genetics and race would today represent a minority view among scientists and scholars in the fields most related to the study of these topics. Shockley believed that present-day

IQ tests are valid measures of intelligence, are racially unbiased predictors of educational achievement and job performance, that individual differences in IQ have substantial heritability and both genetic and environmental factors are involved in the difference between the IQ distributions in the black and white populations. On each one of these points, a majority of experts in the relevant fields express the very same opinions, as shown in a recent large-scale survey (M. Snyderman & S.L. Rothman *The IQ Controversy: The Media and Public Policy*. Transaction Books, 1988).

ARTHUR R. JENSEN

University of California,
Berkeley, California 94720, USA

Current opinion

SIR—Peter Coles (*Nature* 343, 399; 1990) is pessimistic about the present practicability of electric cars when he says, correctly, that to replace the small family petrol-engined car a daily maximum range of around 500 km would be needed. The 100 km now achievable would, however, be quite adequate for the second car of many two-car families, used for shopping and taking the children to school. With good advertising, and perhaps reduced car tax, the demand should cover as many cars as the industry can produce during the period needed for a major improvement in performance.

J. M. FREMLIN

46 Vernon Road,
Edgbaston,
Birmingham B16 9SH, UK

Inverse tomograph

SIR—Daedalus has suggested (*Nature* 343, 516; 1990) that construction of an inverse tomograph could sharpen the accuracy of radiotherapy. You may be interested to know that the details of how to compute the intensities and directions of the beams were worked out here during 1989 using an optimization method which exactly mirrors the X-ray tomography problem¹. This idea was also in the mind of Professor Anders Brahme at Sweden's Karolinska Institute².

But there are snags. To do the job properly, some beam intensities would need to be negative. If Daedalus could design a machine that could achieve this, we should like to see it. A machine that could accurately deliver positive multiple-element beams remains the radiotherapy physicist's dream.

S. WEBB

Joint Department of Physics,
Royal Marsden Hospital,
Sutton,
Surrey SM2 5PT, UK

1. Webb, S. *Phys. Med. Biol.* 34, 1349–1370 (1989).

2. Brahme, A. *Radiother. Oncol.* 12, 129–140 (1988).

Science and scientific attitudes

S. Chandrasekhar

On Apollonius, Kepler and Einstein, Newton and Shakespeare, and Madonna and Mrs Pelham.

ONE of the matters to which I have devoted some thought in recent years concerns one's motivations in the pursuit of science; and some of my thoughts on the subject are included in the collection of my lectures, *Truth and Beauty: Aesthetics and Motivations in Science* (University of Chicago Press, 1987). In this essay, I wish to consider some related questions: (1) the aspect of nature that is in some sense the most incomprehensible; (2) the goals that one strives for in the pursuit of knowledge; and finally (3) the sources of one's satisfaction in such pursuits.

While these questions may seem disparate, it will appear in the discussion that follows that they are in fact related, even if only tangentially.

1

In the context of the first of the topics I have enunciated, I shall consider the following statement of Einstein:

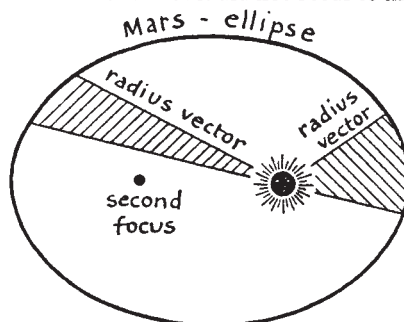
The most incomprehensible fact about Nature is that it is comprehensible.

This expresses a profound truth and it is echoed in the writings of other great men of science. Thus, Eugene Wigner has written of the two miracles: "the miracle of the existence of laws of Nature and the miracle of the capacity of the human mind to divine them"; and he has also written about "the unreasonable effectiveness of mathematics in the understanding of Nature". And Schrödinger considers that this latter capacity of the human mind to divine nature's laws may well be beyond human understanding.

Let me elaborate on these statements with two illustrative examples.

Geometrical curves, their beauty and their forms, held fascination for the Greeks; they are enshrined in their sculptures and their architecture. The straight line and the circle dominate Euclid; and the perfection of the circle was sacrosanct: it underlies Aristotle's conception of celestial bodies and it provides the basis for Greek astronomy from Eudoxus to Ptolemy. But the Greeks did not confine their studies to straight lines and circles only. They sought curves that would encompass the geometrical properties of straight lines and circles in a more beautiful and harmonious synthesis; and they discovered this harmony in the curves obtained as sections in a cone: the ellipse

and the hyperbola. Their curiosity with respect to these curves was not motivated by any physical fact they discerned. And yet, in the second half of the third century BC, Apollonius of Perga wrote eight monumental volumes devoted to these curves. Apollonius is eloquent about their geometrical properties; he describes them as miraculous. But it did not occur to him,



Kepler on his law of areas (the radius vector describes equal areas in equal intervals of time): "Since I was aware that there exists an infinite number of points on the orbit and accordingly an infinite number of distances [from the Sun] the idea occurred to me that the sum of these distances is contained in the area of the orbit. For I remembered that in the same manner Archimedes too divided the area of a circle into an infinite number of triangles".

Kepler on his law concerning the elliptical nature of the orbits of the planets: "Why should I mince my words? The truth of Nature, which I had rejected and chased away, returned by stealth through the backdoor, disguising itself to be accepted . . . I thought and searched, until I went nearly mad, for a reason, why the planet preferred an elliptical orbit . . .".

or to the other Greek mathematicians, that the curves, which they studied so earnestly for their intrinsic beauty, had any relevance to the real physical world.

Yet some eighteen centuries later, when Kepler was analysing the orbits of the planets on the copernican system, he discovered that the very curves that the Greek mathematicians had studied for their intrinsic mathematical beauty were exactly those needed to represent the orbits of the planets. Commenting on this discovery of Kepler, Einstein has written:

Our admiration for Kepler is transcended only by our admiration and reverence for the mysterious harmony of Nature in which we find ourselves. Already in antiquity, man had devised curves exhibiting the simplest forms of regularity. Among these, next to the straight

line and the circle, the most important were the ellipse and the hyperbola. We see that the last two are embodied — at least very nearly so — in the orbits of heavenly bodies.

Einstein continues:

It seems that the human mind has first to construct forms, independently, before we can find them in things. Kepler's marvellous achievement is a particularly fine example of the fact that knowledge cannot spring from experience alone but only from a comparison of the inventions of the intellect with the facts of observation.

Let me repeat the crucial part of this remarkable statement: "The human mind has first to construct forms, independently, before we can find them in things".

I should like to illustrate this same aspect in Einstein's own development of his general theory of relativity, a theory that has been described as "the supreme example of speculative thought".

More than one hundred and thirty years ago in a celebrated lecture, Bernhard Riemann, analysing the foundations of geometry, came to the view that space is not a lot of points close together but a lot of distances interlocked; and further, that these metrical properties of space are causally linked with its matter-content. He states:

Therefore, either the reality on which our space is based must form a discrete manifold or else the reason for the metrical relationships must be sought for, externally, in the binding forces acting upon it.

With this conception of the possible role of geometry in describing the physical world, Riemann developed a system of geometry more general than Euclid's, a system which has come to be known as riemannian geometry.

Some sixty years later, Einstein was absorbed in trying to understand a miracle that underlies the newtonian theory of gravity and an inconsistency in it which he perceived. The miracle is that all bodies are equally accelerated in a gravitational field, as manifested by Galileo's well-known demonstration from the leaning tower at Pisa. It requires the inertial mass and weight to be identical — an enigmatic fact well established but not really understood before Einstein. And the inconsistency that Einstein perceived is that Newton postulated that the action of gravity at a distance is instantaneous — contrary to Einstein's own prohibition that no signal can be propagated with a velocity exceed-

This essay is the substance of an address delivered at Roorkee University, India, on 28 November 1989.

ing that of light. In resolving these paradoxes, Einstein found that Riemann's geometry provided him with a framework ideally suited for developing his general theory of relativity; not unlike Kepler who found in the geometrical works of Apollonius the framework for his understanding of planetary motions.

And Einstein's theory, developed on what appeared to many of his contemporaries as a very frugal basis on which to found a physical theory, has revealed a richness of content that continues to baffle the imagination.

2

Accepting our 'inability' to understand the 'reasonables' of Nature, what are we to understand by the common phrase 'The pursuit of knowledge' as it applies to science?

'Pursuit' has the common meaning of a chase in hunting; and, characteristic of our times, we are also familiar with the word in the combination 'pursuit-plane'. Is this the association which the word 'pursuit' evokes intended? And if it is, are we to conclude that knowledge, like the fox in a chase, or the enemy plane in a hot pursuit, is something whose existence we know of in advance and our pursuit is to aim for it. Of course, some aspects of what we include under knowledge are in this category. Thus, to unearth the fossil remains of creatures of long past, or the relics of bygone civilizations, to scale the highest peaks, or to fathom the deepest oceans, are all endeavours of high human aspiration.

But one may still ask: is knowledge, then, something which we seek to attain in the same spirit as the mountaineer who aspires to ascend the Everest 'because it is there'? If that is the case, what are we to understand when we are told that research is a quest after the unknown to chart territories whose existence we may not even be aware of when we started? Kepler did not know, when he began his long and arduous analysis of the centuries of accumulated observations on planetary motions, that buried in that mass of details were the simple laws he discovered. Neither did Newton know, before he saw the apple fall, that Kepler's laws could be understood so simply in terms of his laws of motion and of gravitation.

Perhaps I shall be accused of quibbling. For one may, in truth, say that in the pursuit of scientific knowledge — if I may be pardoned for that combination of words — while one does not aim for something which is material and concrete, one does aim for the enlargement of that order and that harmony which are the key signatures of nature. Indeed, for a scientist, the order, the harmony, the uniformity and the universality of the laws of nature are as real as the peak of the Everest is for the mountaineer. And so it is in the

other branches of abstract knowledge.

But is that all there is to our quest for knowledge? Do we wish, for example, to quantify new knowledge by the extent to which others can share in it and still others can make use of it for human delight or for human welfare? And if that is our wish, what value do we attach to the refinement of one's own perception and to the enlargement of one's own vision? Is there no real content to Wordsworth's famous lines on Newton:

The marble index of a mind for ever
Voyaging through strange seas of Thought
alone.

Indeed, there is ample evidence that the very greatest artists, in their ennobled maturity, withdraw into themselves. Here, for example, is T.S. Eliot commenting on the late plays of Shakespeare.

It seems to me to correspond to some law of nature that the work of a man like Shakespeare, whose development in the course of his career was so amazing, that it should reach, as in *Hamlet*, the point at which it can touch the imagination and feeling of the maximum number of people to the greatest possible depth and that, thereafter, like a comet which has approached the earth and continued away on its course, he should gradually recede from view until he tends to disappear into his private mystery.

What T.S. Eliot has said of Shakespeare applies equally to Beethoven. In his late compositions and especially in his late quartets, Beethoven was indeed "Voyaging through strange seas of Thought alone" — voyaging, in fact, to enlarge his own personal vision.

This attempt by the greatest minds to enlarge their personal vision is, I believe, manifest also in the remoteness and in the laminated and glacial style of Newton's *Principia*. The lasting value of the *Principia* lies as much in Newton's vision of the Universe, as in the surpassing quality of his discoveries which it summarizes and organizes.

3

And finally I wish to consider the sources of one's satisfaction in scholarly endeavours.

Perhaps I should first eliminate, *ex cathedra*, that the rewards of scholarship consist of celebrity and of public honours. The supreme Magisters of Castalia, in Hermann Hesse's *Glass Bead Game*, had learned to renounce them. I suppose that one does renounce them in the end — at least, one feels that one ought to transcend them. But the matter is not as simple as that. None of us are so immune to human sensibilities that we are not all, to some extent, sensitive to the approval of our colleagues whom we respect. And I am sure that all of us hope, each in his own way, that posterity will assign to us our due and humble places so long as we persevere at the limits of our capabilities.

But posterity can be harsh. Here, for

example, is John Ruskin's criticism of the English painter Sir Joshua Reynolds:

Why did not Sir Joshua — or could not — or would not Sir Joshua — paint Madonnas? . . . Yet! While we acknowledge the discretion and simple heartedness of these men . . . we have to remember . . . that amiable discretion is not the highest virtue, nor to please the frivolous, the best success . . . There is probably some strange weakness in the painter, and some fatal error in the age, when in thinking over the examples of their greatest work, for some type of culminating loveliness, or veracity, we remember no expression either of religion or heroism, and instead of reverently naming a Madonna di San Sisto, can only whisper modestly, "Mrs. Pelham feeding chickens."

If one recognizes that one can never come to painting a Madonna, what, then, are the satisfactions and the rewards? I suppose that one must count them in those brief moments of sudden insight which occur to one on rare occasions. One may never come to painting a Madonna. But, perhaps, in capturing on canvas, the rugged lines in the face of Mrs Pelham, etched by the toils of her life, the painter may have experienced a sudden insight into the sadness of the human condition which he may cherish all his life. And so it is in all other walks of creative effort.

While one may grant that the rare moments of illumination one experiences during the course of one's life are the precious rewards, one is still troubled; for, one may ask: is he condemned to live only in the memory of his past moments of illumination? The answer, it seems to me, is that, if one is not to be so condemned, one must seek in one's enlarged and refined perception a source for quiet contemplation. In other words, "Voyage through strange seas of Thought alone". But even this may not always be possible, as Hermann Hesse poignantly describes in his portraiture of Joseph Knecht in the *Glass Bead Game*.

The *Glass Bead Game* is the wondrous tale of one who, in his youth, aspired to the ideals of Castalia and who, in time, rose to become the supreme Magister Ludi. But in the end, in deep doubts, he renounces his office, exiles himself from Castalia, only to be drowned accidentally in a Swiss lake.

Let me try to answer directly the question why one, fully aware of one's inherent and often insurmountable limitations, nevertheless devotes oneself to scholarship and to a life of constant endeavour, many failures and fewer successes? The answer is, as T.S. Eliot has given in his "Confidential Clerk":

It is strange, isn't it,
That a man should have a consuming passion
To do something for which he lacks the
capacity?

S. Chandrasekhar is in the Laboratory for Astrophysics and Space Research, The University of Chicago, 933 East 56th Street, Chicago, Illinois 60637, USA.

Towards the electronic journal?

Recent developments in optical disk technology promise electronic access to whole libraries of scientific journals, but the consequences for the ways in which journals operate are far from clear.

In major technical innovation, hopes and ambitions lead reality by years, even decades. So much is common experience. From the late 1950s onwards, people have been making public speeches (and organizing conferences at which to make them) on the theme that computer technology would change the world, but two decades went by before it did so. Much the same happened with air transport, technically possible since early this century, but a means of mass transport only after more than half a century. Why should there be these lags? And how is it possible to tell when they are about to end?

Two sets of considerations arise. First, the enthusiasms of the prophets of innovation customarily outstrip what technology will allow. (Arthur C. Clarke, who sang the praises of communications satellites in the 1950s, was lucky that he had to wait a mere decade to see something of the reality.) Second, and probably more important, the potential beneficiaries of innovations tend to stick to their old-fashioned ways in the face of all the evidence that change would be beneficial.

The same two influences explain why the late J. D. Bernal's dream that the scientific literature would be organized so that its elements could be directed exclusively towards those with particular interests in them has been on the shelf for more than four decades: the technology has been inadequate and the users indifferent.

But revolution is now in the air. There has been no single change of attitudes, nor single improvement of technology, to account for what is happening. Rather, a dozen different changes, and the prospect of more ahead, has given the notion of the electronic journal (more strictly, journal-complex) the promise of reality.

Technology is the *sine qua non*, and has markedly improved. This page of *Nature* will hold roughly 7,250 bytes of information when it is complete. (The figure is not exact because printed characters differ in width.) The whole of a weekly issue's text would fit comfortably into a megabyte of computer storage. But it is only in the past two years or so that information storage of that magnitude could be had at a reasonable price. And even now, storing a year's issues of this journal in a conventional computer store would be an extravagant proposition; even from a mail-order address, a 50 megabyte hard disk would cost the best part of \$1,000.

Optical disks have changed all that.

Devices not much more sophisticated than those sold in audio-record retail shops will hold 500 megabytes, enough for a decade of *Nature*. True, these disks are 'read-only' disks, but who wants an archive that can be rewritten by its users? The same device could be used to put out weekly issues of 500 different weekly publications of roughly the same extent as this. The problem of storing information has been solved, it may appear.

This is the present inspiration of the *Adonis* project by which a consortium of European publishers (Blackwells, Elsevier, Pergamon and Springer) plans to start next year distributing to subscribers a weekly issue of an optical disk carrying the stored images of pages from the current issues of more than 400 scientific journals. Already, the plan embodies elements of Bernal's vision of a user-directed system of document supply. The first customers will be the libraries at pharmaceutical companies' research centres; the journals to be distributed will be chosen appropriately.

Adonis differs from a legion of earlier schemes in the assurance one has that it will function well. The consortium has just emerged from a two-year trial with just over 200 journals, distributed on optical disks to a handful of libraries around the world, which has tested the latest somewhat technically extravagant version of the technology. The image of each page of a journal is represented in facsimile, not by character-coding, as the stream of bits of information arising when the image is scanned at specified resolution. 'Bit-mapping' is the word.

Two obvious advantages are that the system will reproduce pages from particular journals with exactly the appearance they have in printed form, and that it will handle illustrations, even photographs. In practice, it is planned to scan text with a horizontal resolution of 300 picture elements per inch and a vertical resolution half as fine; that should deal well with characters a third as high as those on this page. The vertical resolution of the scanning of photographic illustrations will be twice as fine as that for text.

Fears that the simple-minded scanning of printed pages will use enormous quantities of storage have been dispelled by clever software. The system is smart enough not to store the results of scanning with white spaces, but allows the decoding system to reconstruct them. No doubt it is

only a matter of time before character-coding and bit-mapping are mixed in a bastard system, but meanwhile it is feasible to include more than 5,000 typical pages from scientific journals on a single disk with some assurance that their contents can be accurately recovered as computer images or even as hard copy.

But what value is an optical disk embodying the equivalent of more than a year's supply of *Nature* if its contents are not self-evident? None at all, of course. So *Adonis* has been working at the design of an index that will allow users not merely to reconstruct its contents but also to allow individuals to select the items likely to be of interest in their research. 'Research profile' is the buzz-word.

Will it work? *Adonis* is naturally confident. Technically, there seems little doubt. But how will potential customers respond? After two decades of thumbing through *Current Contents*, marking up (for reprint requests) an appropriate version as if it were a racecard, and similar experience of the databases of the US Library of Congress and the American Chemical Society, one senses that there is a generation of researchers whose need of information is menu-driven. If (as *Adonis* hopes) a researcher's profile is instantly required by laser printout, reactions may be much more favourable than in the past. Bernal's dream may be coming true.

But will it also be a nightmare? Technically, it will be a jungle; *Adonis* is not the only commercial enterprise in the field, while it is improbable that non-profit organizations such as the American Chemical Society and the US Library of Congress (which has been patiently experimenting for the past five years) will allow themselves to be left behind. Then there lies ahead the whole multimedia technology, whose broadband communication systems may supersede disks that must be physically transferred from place to place.

Intellectually, it is not a pleasing prospect that the distinctive identities of journals will inevitably — not quickly, but in the course of time — be subsumed in the identity of their carrier, the purveyor of the disks. Amalgamation into megajournals is a likely outcome, but what will authors think of that? Browsing may not die off, but will have to be pre-planned. That is why it must be hoped that there will always be some ink-on-paper journals.

John Maddox

The Chinese side of the story

Alison S. Brooks and Bernard Wood

THE first human fossil remains recovered in China were found at Sjava-osso-gol (Salawusu), in the Ordos region of Mongolia in 1922. But it was the discovery of the first cranium of 'Peking Man', at Choukoutien (Zhoukoudian) in 1929, that turned attention to China and marked the beginning of the series of important fossil and archaeological discoveries that continues to the present time. These discoveries come from all three stages of the Pleistocene, as well as the Holocene, and the remains from Gongwangling at Lantian may be among the earliest fossil hominid evidence outside Africa. The Institute of Vertebrate Paleontology and Paleoanthropology of the Academia Sinica has played a pivotal part in the discovery, excavation and analysis of much of this evidence, and it recently convened a symposium to mark the sixtieth anniversary of the original find*. The occasion was especially valuable because the results of the work of many Chinese researchers dispersed throughout the country are not always widely disseminated within China, let alone outside it.

Most prehistorians share in the growing realization that Europe was peripheral to many of the main events in human prehistory; the australopithecines were almost certainly confined to Africa and the emergence of our own species, *Homo*, can at present be traced only through the African fossil evidence. There is, however, a complementary realization that Asia, and particularly China, may hold crucial evidence for the interpretation of later stages in the physical evolution and cultural elaboration of *Homo*. These stages include the movement of *Homo*, probably as *Homo erectus*, into the higher latitudes, the emergence of archaic *Homo sapiens* and, later, of anatomically modern *Homo sapiens*, together with the nexus of behavioural changes that drove or accompanied these evolutionary developments and resulted in modern human cultures.

The main emphasis of the contributions devoted to the hominid fossil evidence was their chronometric dating. The collection of fossil hominids from locality 1 at the Zhoukoudian site comprises 14 crania and cranial fragments, 16 mandibles, teeth and limb bones, all now attributed to *Homo erectus*. Only those fossils found since the 1950s survive. The bulk of the collection was lost in mysterious circumstances during the Second World War. So far all attempts to locate the missing specimens have proved unsuccessful, but

so accurate are the casts that were prepared in the United States, and so meticulous were the observations, photographs, radiographs and drawings made by Franz Weidenreich that the loss of the originals was not the disaster it might have been. The casts and fossils can potentially furnish information about development and sexual variation within an early hominid taxon, as well as providing a useful tool to investigate whether the mode of hominid evolution conforms to either of the gradualistic or punctuational models that have been proposed.

Until recently, progress with these research programmes was hampered by the lack of sound and detailed dating evidence. But new fission track, electron spin resonance (ESR), thermoluminescence, and uranium series dates (T. M. Chen, Peking University; Y. M. Guo *et al.*, Institute of Atomic Energy; and S. S. Zhao, Institute of Geology) for locality 1 are producing internally consistent results. The hominid fossils were concentrated in layers 8–10, with one particularly progressive, or derived, skull (V) located in level 3. Layers 9 and 10 date from around 460 kyr, whereas the results for layers 1–2 indicate that the top of the sequence is dated to about 230 kyr. Concordance between different dating techniques is particularly impressive for level 4 ash horizons using the fission track (306 kyr), thermoluminescence (292±26 kyr, 312±28 kyr), U series (300±4.1–3.0 kyr) and ESR (293.13 kyr) techniques. This evidence suggests that *Homo erectus* was living, without evidence of significant morphological change, in and around Zhoukoudian for more than 200 kyr. In

fission track dating it is assumed that the crystals were completely annealed at high temperatures during deposition, so the agreement between the fission track dates and those obtained by other techniques also implies the use of fire at the site.

Somewhat more of a problem are the dates for several later sites, for which U series is the only dating technique used¹. Even if different U-series ratios (²³⁰Th/²³⁴U and ²³¹Pa/²³⁵U) yield similar ages, secondary uptake or leaching of uranium cannot be ruled out. If correct, the dates suggest temporal overlap in China between the latest *erectus* fossils at Zhoukoudian and the archaic *Homo sapiens* remains from locality A at Jinniushan, in Liaoning province to the northeast of Zhoukoudian. At the latter site, technically difficult excavations were carried out in a series of fissure fills (Z. Lü, Peking University), leading to the recognition of seven layers at locality A itself. The male skeleton found in 1984 in the lowermost layer (layer 7), underlies a horizon with at least six U-series dates on dentine centring around 220–314 kyr, with an average of 263 ± 30 kyr. The U-series age of 67 kyr for the anatomically modern Liujiang fossil from a cave in the Guangxi region of South China is surprisingly early, for most other essentially modern-looking remains from Eurasia date from about 40 kyr, or less. But the early date is consistent with the associated late Middle/early Upper Pleistocene faunal remains, as well as with the retained archaic features in the fossil itself.

In an effort to overcome political barriers to scientific collaboration, the conference included the first and historic meeting of colleagues from North and South Korea, who together made up half of the foreign delegation. Papers they contributed included a description of the archaic features of a new North Korean

Chronometric dating of Chinese hominid fossil localities*

Species	Locality	Age in kyr
<i>H. erectus</i>	Yuanmou	~ 1700 (M), 500–600 (M)
	Gongwangling	750–850 (M), 1,000 (M), 1,150 (M)
	Chenjiawo	650 (M)
	Zhoukoudian	230–500 (F,U,A,T,E)
	Hexian	150–200 (U) <200 (T)
<i>H. sapiens</i>	Dali	180–230 (U)
	Jinniushan	210–300 (U)
	Changyang	170–220 (U)
	Dingcun	160–210 (U)
	Tonzi	172–192 (U), 214 (U)
	Xindong, ZKD	135–175 (U), 100–200 (U)
	Maba	119–140 (U)
	Xujiayao	83–131 (U)
	Jiande	90–117 (U)
	Liujiang	67 (U)
<i>H. s. sapiens</i>	Salawusu	37–50 (U), >35 (C)
	Ziyang	7 (C), 36–39 (C)
	Upper Cave ZKD	~10–20 (C)

* From refs 1–3, and Wu Xinzhì, Conference paper. Dating by: M, palaeomagnetism; F, fission track; U, uranium series; A, amino-acid racemization; T, thermoluminescence; E, electron spin resonance; C, radiocarbon.

* International Symposium on Palaeoanthropology Commemorating the 60th Anniversary of the Discovery of the First Skull of Peking Man, Fangshan, China, 19–24 October, 1989.

fossil hominid from Yongok-ni (U. J. Jan), and a demonstration of the co-existence at Chon-gok-ni of both large- and small-tool traditions in a single site (K. Bae, Seoul). The large tools from Chon-gok-ni are more like the true hand-axes than most material from east Asia, but the site, like the 'Lower Palaeolithic' industries of Kota Tampan in Malaysia (Z. Majid) and the Patjitanian of Java, seems to be Upper Pleistocene in age.

The conference re-established the importance of China in the overall scheme of later human evolution, and showed that this region is best understood on its own terms rather than by comparing it with western Eurasia. Our understanding of the Chinese palaeolithic record has been hindered by the application of western European parameters of technological development. The widespread use of expedient technologies on poor raw materials such as quartz, the persistence of crude bipolar cores, chopping tools and other 'Lower Palaeolithic' forms into the later Pleistocene, and the absence of well

developed blade industries in many regions, all suggest a pattern more common to areas of central and west Africa than to western Europe. We can hope that when 'Peking Man' celebrates his diamond anniversary, the dramatic expansion of research in East Asia will have shed further light on the precise nature of the apparently unique human adaptations that characterize this region, as well as on the relationships between East Asia and Africa, during the emergence of our own species. □

Alison S. Brooks is in the Department of Anthropology, George Washington University, Washington DC 20052, USA. Bernard Wood is in the Department of Human Anatomy and Cell Biology, Liverpool University, PO Box 147, Liverpool L69 3BX, UK.

1. Chen, T. & Yuan, S. *Archaeometry* **30**, 59–76 (1988).
2. Wu, R. & Olsen, J. W. *Palaeoanthropology and Palaeolithic Archaeology in the People's Republic of China* (Academic, Orlando, 1985).
3. Wu, R., Wu, X. & Zhang, S. *Early Humankind in China* (Institute of Vertebrate Paleontology and Paleoanthropology, Beijing, 1989) (in Chinese).

QUANTUM ELECTRODYNAMICS

Photons faster than light?

Stephen M. Barnett

ONE of the cornerstones of twentieth-century physics is the limiting velocity of the speed of light, c . Every undergraduate knows that no particle can travel and no signal can propagate at a velocity greater than c . This comfortable and well established phenomenon has now been challenged by Scharnhorst¹ and Barton². They have uncovered a quantum electrodynamic phenomenon that implies the possible existence of faster-than-light photons.

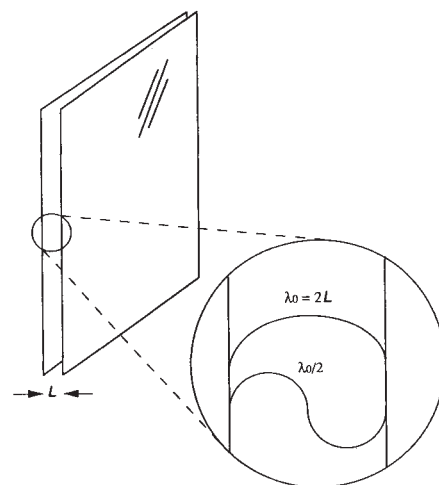
The source of this surprising development is the mysterious nature of the vacuum as revealed by quantum electrodynamics. It is not an empty nothing but contains randomly fluctuating electromagnetic fields and virtual electron-positron pairs with an infinite zero-point energy. Light propagating through space interacts with the vacuum fields, and observable properties, including the speed of light, are in part determined by this interaction. The effect on the speed of light is due to the absorption of photons to form virtual electron-positron pairs followed by the re-emission of the photon. The particle pairs are very short lived because of the large mismatch between the energy of a visible photon and the rest mass of an electron and a positron. Nevertheless, this process makes a contribution to the observed permittivity of the vacuum (ϵ_0) and therefore to the speed of light.

The role of the virtual particle pairs in determining the permittivity of the

vacuum is analogous to that of atoms or molecules in determining the relative permittivity of a dielectric material. The light propagating in the material can be absorbed, but in a transparent medium there are no strongly absorbing resonances and the atoms remain in their excited states for only a very short time before re-emitting the light. This absorption and re-emission is responsible for the refractive index of the material and results in the well known reduction of the speed of light.

Scharnhorst and Barton suggest that a modification of the vacuum can produce a change in its permittivity with a resulting change in the speed of light. In particular, the suppression of light-scattering from the virtual electron-positron pairs should cause an increase in the speed of light. One way to change the vacuum is to place two large conducting plates parallel to each other, so as to form a cavity or waveguide (see figure). The vacuum field inside the cavity has to be modified so that the electromagnetic boundary conditions are satisfied. The principal result of this modification is the removal of the zero-point energy due to the suppression of vacuum modes with wavelengths longer than the waveguide-cutoff (λ_0). It is this removal of free-space vacuum modes that suppresses the scattering of the light by virtual particle pairs and in theory produces the increase in the speed of light.

The predicted increase in the speed of light is extremely small and only occurs



The vacuum field is modified in a waveguide and electromagnetic modes with wavelengths greater than λ_0 are suppressed.

for light propagating perpendicular to the plates. For a plate separation of $1 \mu\text{m}$ the fractional change in the speed of light is about one part in 10^{36} . This effect is so small that it is far beyond the reach of present experiments, but its significance as a point of principle should not be underestimated. It is worth emphasizing that the increase in the speed of light is a genuine prediction for the propagation velocity. Optical phase and even group velocities can sometimes exceed c , but dispersion always ensures that the signal velocity is less than or equal to c . In the parallel-plate system described by Scharnhorst and Barton, the dispersive effects are much weaker still than those associated with the apparent increase in the velocity of light; the phase, group and signal velocities will therefore all increase by about the same amount.

The reduction of the vacuum energy density in the region between two parallel conducting and neutral plates has other consequences, most notably the famous Casimir effect. If the plates are moved towards each other then the vacuum energy drops. This means that there is an attractive potential between the plates, the origin of which is purely quantum electrodynamic in nature. The resulting force pulling the plates together has been measured and for a plate separation of $0.1 \mu\text{m}$ has a value of about 13 newtons per square metre of the plate surfaces.

The vacuum is certainly a most mysterious and elusive object that makes itself known by only the most indirect of hints. The suggestion that value of the speed of light is determined by its structure is worthy of serious investigation by theoretical physicists. □

Stephen M. Barnett is in the Department of Engineering Science, University of Oxford, Parks Road, Oxford OX1 3PJ, UK.

1. Scharnhorst, K. *Phys. Lett. B* **236**, 354–359 (1990).
2. Barton, G. *Phys. Lett. B* (in the press).

Simplifying complexity

Alan R. Bishop

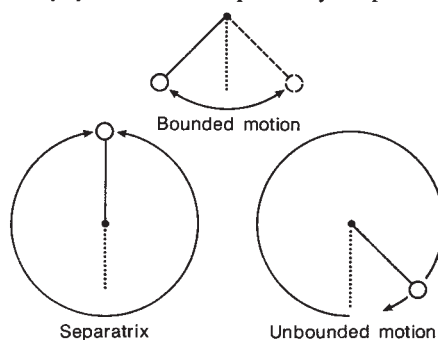
WE are in a period of great enthusiasm for 'complex systems', which in a mere decade has grown to include almost every discipline and scale in the natural sciences (for example, see ref. 1) from astrophysics to biology, and even the less natural sciences such as the stock market and arms control. In this quest for common themes and paradigms of complex reality, it is fitting to be reminded that: first, the most fundamental underlying theme is the classical dynamics of mathematically non-integrable systems; second, we still have only a primitive understanding of even simple-looking few-body systems; and third, nontrivial few-body problems in nature have rarely been brought under the control of the experimentalist to the stage of testing our understanding of chaotic dynamics. The experiment of Brewer and colleagues² and their theoretical analysis on page 305 of this issue³ provide a salutary lesson and example: the seemingly simple system of two ions confined by a radio-frequency field and cooled by radiation pressure exhibits a transition from periodic to chaotic motion, described by an unexpectedly rich route to chaos and a 'strange attractor'. The route supplements more familiar ones such as period doubling (found for example in population dynamics) and intermittency (as seen in the onset of turbulence).

According to the analysis by Brewer *et al.* of the ions' dynamics, which leads to an accurate description, the motion is mostly regular, but occasional hard collisions between the ions result in transient instabilities; these 'reset' the initial conditions in an uncorrelated way. A similar situation prevails for other two-body potentials with a hard-core repulsion and soft attraction, but the experimental accessibility here is unusual: many other controlled experiments have the complication of being either intrinsically quantum mechanical (for example, the hydrogen atom in a radio-frequency electric field or a strong magnetic field) or macroscopic (as in fluids or plasmas).

The issue brought into focus is how well we understand classical mechanics, which is unfortunately not as clear as we would wish. The elegant geometric and algebraic mathematics of the last century were largely displaced by the advent of quantum mechanics, despite our continuing inability to unravel the full subtleties of quantizing nonintegrable classical systems. Most textbooks have little to say about any more than integrable equations, and it is only in the past few years that students of all ages have begun to rediscover classical mechanics and a host of old but unsolved problems⁴.

The concepts of integrability and its absence are crucial: the definition of integrability is intensely debated. The task of delineating classes of few-body systems that are integrable within standard definitions has involved elegant formalisms from many branches of mathematics (see, for example, ref. 5). Such integrable hamiltonian (that is, friction- or dissipation-free) systems are of course a highly unrepresentative subset of all hamiltonians: even the motion of heavenly bodies is now thought to be weakly chaotic—it is to be hoped on a heavenly timescale. Nevertheless, these integrable systems are frequently used as a starting point to understand the effects of structural or additive perturbations.

The motion of the particles in a few-body system can be depicted by the path of



Pendulum acting nonlinearly under gravity, showing separatrix between bounded and unbounded motions.

a single point on a multi-dimensional graph whose coordinates indicate the position and momentum of each particle at any particular moment. For a hamiltonian system, in which the total energy is unchanging, this path is confined to a surface of constant energy in this 'phase space'. If the equations are integrable, there are many other conserved quantities which further confine the possible trajectories to a lower-dimensional toroidal surface.

Adding perturbations can make the system less integrable, for which the paths through phase space are more complicated and irregular. For example, 'resonances' can occur as described by Chirikov⁶, where natural frequencies lock at various commensurate ratios. If the paths of these resonances overlap in phase space, the behaviour can diffuse to distant parts — as may happen, for example, in the ionization of highly excited hydrogen. The manner in which perturbations can change the regular motion, depicted by closed paths, of integrable systems into this 'ergodic' diffusion throughout the allowed phase space is one of the great challenges of mathematical physics.

Hamiltonian systems (integrable or not) have the property of preserving areas, that is, staying within a given volume, in phase space. This is to be contrasted with the equally prevalent types of dynamics, which include both energy input (for example, through an external field, as in the case described by Brewer *et al.*²) and dissipation, although dissipation raises some fundamental difficulties for a quantum-mechanical description. The paths in phase space tend to become confined to a finite number of basins of attraction; that is, the motion collapses onto a subspace of the full phase space of the system. They may be simple fixed points or periodic or quasi-periodic (with two or more incommensurate frequencies) attractors — or, in the parlance of modern dynamical systems⁴, the attractors may be 'strange'. These objects are characterized by motions on a 'fractal' or noninteger-dimensional component of the full space and by an extreme sensitivity to initial conditions, as measured by Lyapunov exponents and described by Brewer *et al.*³. The dynamics on the phase space accessible to the strange attractor may seem completely irregular, but it is entirely deterministic. This is different from the effect of external white noise, which allows the dynamics to explore the whole of phase space (that is, to become ergodic). The approach to chaos and motion on a strange attractor from regular dynamics occurs as parameters are changed (such as the frequency or strength of a radio-frequency field). The nature of this transition has begun to be classified for simple low-dimensional dynamical systems. Feigenbaum⁷ and others have identified a period-doubling sequence; other systems show intermittent bursts of chaos separating quiescent periods. A full classification of routes to chaos still eludes us, but an important step has been to convert the nonlinear differential equations that describe the motion into 'maps' of the system's evolution over discrete time steps (by sampling the dynamics periodically or stroboscopically, for example). This was suggested by Poincaré in the last century⁸: certain maps derive from classes of equations and display 'universal' routes to chaos^{4,7}.

It may be helpful to think of a concrete single-particle example, namely the familiar pendulum acting under gravity. The pendulum in its full nonlinear glory has a 'separatrix' between bounded and unbounded motions — that is, between finite periodic oscillations and full 360° rotations (see figure). The separatrix motion corresponds to an oscillation that sweeps through the entire 360°; it is an infinite-period solution taking an infinite time for the pendulum initially at rest in the unstable vertical position to rotate by 360° back to this position. If the pendulum is now driven, for example by applying a

force via a radio-frequency field, the motion will pass through a deeply complex sequence of attractors as the driving strength and initial conditions are varied. The Soviet school has developed tests (see ref. 9, for example) for the onset of non-trivial dynamics in terms of the effects of small perturbations on the separatrix solutions of the unperturbed dynamics — sufficient perturbation allows the pendulum to visit the bounded and unbounded motions nonperiodically.

The 'simple' pendulum embodies a truly common ingredient of complex dynamical systems, namely, the competition between timescales that can lead to frequency locking, quasiperiodicity and eventually chaos. The timescales in the driven (or kicked) pendulum are the driving frequency and the natural free-pendulum frequency. In other problems there may be several competing frequencies. Precisely the same ingredient extends to purely static problems with competing length scales or lattice symmetries — for instance, epitaxial surface physics or quasicrystals. Here the complexity manifests itself as complicated spatial patterns: in formal terms, discrete time maps can be exchanged for discrete spatial maps.

Ultimately we must learn to combine time and space (so generalizing from few- to many-body systems). This brings us back to the full complexity of nature and a whole new set of issues¹⁻⁴: the seeming paradox that chaos is possible in few-body systems, whereas statistical mechanics is

based on the statistical randomness of systems with a very large number of degrees of freedom; that spatially extended, nonlinear systems frequently display strong coherence, pattern formation and low-dimensional chaos quite similar to the few-body dynamics. The resolution here is the beautiful ability of many nonlinear partial differential equations to form long-lived, coherent, particle-like space-time structures, whose dynamics (including chaotic dynamics) can indeed take place in a low-dimensional space. This is distinct from 'turbulence', however, which reflects dynamics in many degrees of freedom, and provides largely virgin territory for future generations of researchers into dynamical systems. □

Alan R. Bishop is in the Center for Nonlinear Studies and Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA.

1. Bishop, A. Gruner, G. & Nicolaenko, B. (eds) *Spatio-temporal coherence and chaos in physical systems*, *Physica* **23D**, 1–340 (1986).
2. Hoffnagle, J., DeVoe, R.G., Reyna, L. & Brewer, R.G. *Phys. Rev. Lett.* **61** 255–258 (1988).
3. Brewer, R.G., Hoffnagle, J., DeVoe, R.G., Reyna, L. & Henshaw, W. *Nature* **344**, 305–309 (1990).
4. Cooper, N. (ed.) *Stanislaw Ulam, Los Alamos Science* **15**, (1987).
5. Ferreira, L.S., Fonseca, A.C. & Streit, L. (eds) *Models and Methods in Few-body Physics, Lecture Notes in Physics*, Vol. 273 (Springer, Berlin, 1987).
6. Chirikov, B.V. *Phys. Rep.* **52**, 278–312 (1979).
7. Feigenbaum, M.J. *Physica* **7D**, 16–40 (1983).
8. Poincaré, H. *Les Methodes Nouvelles de la Mécanique Céleste* (Gauthier-Villars, Paris, 1892).
9. Lichtenberg, A.J. & Leiberman, M.A. *Regular and Stochastic Motion* (Springer, New York, 1983).

PLANT BREEDING

Seeding the bamboo revolution

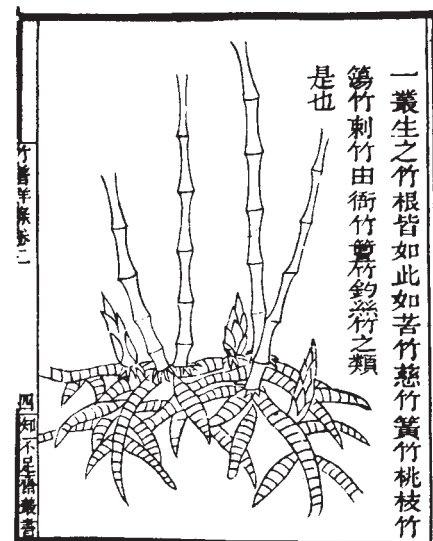
David E. Hanke

It is difficult to exaggerate the importance of bamboo as a structural raw material for most of humankind. Yet throughout the long history of its use, the timber and fibre crop that underpins Eastern culture has never been improved by selective breeding. Its appallingly long generation time puts it way out of reach of any geneticist — how do you maintain a programme of recurrent selection with an interval of 120 years between crosses? A remarkable discovery, reported on page 335 of this issue by Nadgauda, Parasharami and Mascarenhas¹, has broken the deadlock and at a stroke delivered "the most versatile grass"² into the hands of plant breeders. Nadgauda and colleagues found that tissue-cultured shoots from bamboo seedlings on medium supplemented with a cytokinin (N_6 -substituted adenines, active in plant development) and coconut milk flowered in appreciable numbers after only three subcultures. This is extraordinary in that, intact and in soil, seedlings of the two species concerned would with dogged persistence have grown

without flowering for 30 years.

It is by no means the first report of acceleration of the onset of flowering in juvenile material as a result of tissue culture. Loblolly pine³ and the day lily⁴ are notable examples, and in the second case ethene, active in plant development, reversibly promoted the transition from juvenile to mature. For tissue-cultured date-palm shoots, cytokinin again contributed to a reduction in the duration of the juvenile phase from the usual nine years or so to five months⁵. But the effect in bamboo is much the most spectacular, and potentially the most important.

The 500-plus species are scattered throughout warmer parts of the world, but the family achieves its greatest abundance and most impressive luxuriance around the southern and southeastern edge of Asia, from the Indian monsoon region through China and Japan to Korea. Here the larger bamboos throw up towering forests, breaking growth records for the vegetable kingdom. The pointed, sheath-folded shoots emerge from the soil at up



Clumped-root habit in bamboo — new rhizomes accumulate over the years in a tangled pile. Picture taken from Li Khan's *Chu Pu* (around 1299).

to 4 cm per hour, soar to 40 m (130 ft) in some species, and generate a full, feathered canopy in two or three months. It helps of course that the stems, up to 30 cm in diameter, are hollow. Individual stems die and are replaced in their third season by stems from new rhizomes that accumulate over the years on layers of pre-existing rhizomes in a tangled pile a metre or so thick (see figure). Smothering all opposition, large bamboos extend over vast tracts of land in monoculture.

There are bamboos which flower annually, but in almost all the species within the Indian-Asian tropics the exclusively vegetative phase is prolonged for years — 15, 30 (the most usual), 60 or even 120 — before the entire stand, from the tiniest sprout to the mightiest column, flowers as one, sets seed as one, and dies as one. The aftermath is truly apocalyptic as rotting culms collapse in heaps, fires rage and hordes of rats gorge on seed that lies up to 15 cm thick on the ground. Spasmodic mayhem, it seems, is the ultimate outcome of the escalation of evolutionary siege-warfare into which the bamboos and their seed predators became inextricably locked. The only bamboos to survive were those with a vegetative phase prolonged beyond the life-time of the predators whose massed ranks were swollen by the seed crop. The additional vicious circle from which the bamboos could not escape is that the vegetative phase can be prolonged only by doubling its duration; any other increase leaves the first representatives of the new genotype vulnerably out of phase with the majority. In pointing this out, Janzen⁶ also argues persuasively that the fecundity of pigs and chickens arose through bamboo seed predation before these animals became domesticated. Perhaps the West owes more to bamboo than would at first appear, although when the rat is

added to the balance the debt must be at least partly cancelled.

The nature of the plant's accurate internal almanac remains obscure. Timing is perfectly maintained by every part of a fragmented plant, and is superbly buffered so as to compensate for environmental, nutritional and developmental differences between the fragments. There are many reports of bamboos brought to European gardens from India and China flowering in perfect synchrony with the parent stand after intervals of 30 years or so⁶. An important experiment, which does not seem to have been attempted, is to check the time-keeping of plants transferred to the Equator. The results of Nadgauda *et al.* suggest that cytokinins may be involved (perhaps with coconut milk supplying the inositol and cytokinin oxidase inhibitors which promote responses to cytokinin). But it would seem just as likely that the applied cytokinin is acting to reverse the recent and preceding transition from mature flowering to juvenile seedling tissue as it is to advance the subsequent transition from juvenile to mature by 30 years.

In the post-industrial world, the outstanding productivity of the plant and versatility of the material will ensure the global importance of bamboo. Bamboo hay, for example, has four times the protein content of hay from fodder grasses⁷ (the giant panda is no fool), and paper from bamboo, perfected in China in remote times, is much better than newsprint. The new discovery should make possible an explosion of new types, both by genetic improvement within species and in the shape of new hybrids. There is plenty of variation to choose from. Foliage leaves, for example, vary between species from great sheets 4.5 m long and 30 cm wide (on a plant only 3 m high) to hair-like threads. There is even a square-section bamboo, *Bambusa angulata*⁸. Further, bamboos will hybridize when they can be brought to flower at the same time, and hybrids between sugar cane and bamboo were obtained as long ago as 1937 (ref. 9). The discovery by Nadgauda, Parasharami and Mascarenhas opens up endless possibilities for the High Emperor of all the Grasses. □

David E. Hanke is in the Department of Botany, University of Cambridge, Downing Street, Cambridge CB2 3EA, UK.

Cytoskeletal ups and downs

Michael Way and Alan Weeds

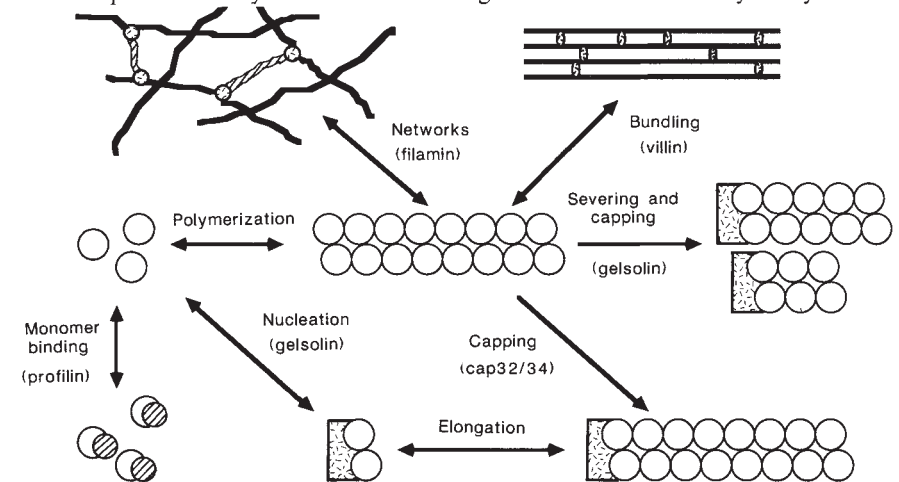
ACTIN-binding proteins provide the driving force for the rearrangements of the actin cytoskeleton in cell motility, division and differentiation. They have been classified according to their properties *in vitro*, but these activities may not fully reveal their cellular functions. Two reports now reveal clear phenotypic differences *in vivo* after genetic manipulation of two different actin-binding proteins. Amatruda *et al.* on page 352 of this issue¹, show loss of normal cytoskeletal organization during budding of yeast cells when the β -subunit of actin-filament-capping protein is deleted. Friederich *et al.* have demonstrated that striking changes occur in the actin cytoskeleton and consequent cellular morphology when a foreign protein, villin, is expressed in fibroblasts. Are *in vitro* and *in vivo* approaches converging to provide a better understanding of cytoskeletal function?

In embryonic carcinoma cells, actin is distributed diffusely in the cytoplasm or concentrated at the cell cortex, reflecting the round shape and low surface-adhesion properties of these cells. Differentiation results in conversion to large flat cells with extensive pseudopods and is accompanied by the formation of well ordered structures containing actin filaments³. Cells that adhere to surfaces and flatten out before locomotion also show a concomitant rearrangement of their cytoskeletons. Because the level of total cellular actin remains relatively constant, cytoskeletal reorganization must be brought about by the interplay of actin regulators.

The figure shows a simple classification of actin-binding proteins. Proteins such as profilin that bind to actin monomers (G actin) seem to be responsible for G-actin buffering in the cytoplasm: they ensure that there is a pool of G-actin monomers for subsequent assembly of actin filaments

(F actin). F-actin-binding proteins are grouped by the nature of their interactions: α -actinins are rod-like dimers which crosslink actin filaments into loose bundles or networks; filamin and gelation factor provide flexible crosslinks for networks as found in the cortical cytoplasm of motile cells. It is emerging from studies of their amino-acid sequences that the actin-binding domains of these crosslinking proteins are similar⁴, suggesting they interact at the same site on the actin molecule. The actin-severing and capping proteins, as typified by gelsolin or severin, seem to be ubiquitous in eukaryotic cells. They sever actin filaments and cap their fast-growing ends in a calcium-dependent manner; they also bind monomers and nucleate polymerization. Internal amino-acid sequence repeats within these proteins suggest that they have evolved from a smaller repeating segment⁵. Villin is also related to gelsolin, but contains an additional actin-binding domain so that, in the absence of calcium, it bundles filaments. A further class of capping proteins, typified by Cap 32/34 (ref. 6), has no sequence similarity with gelsolin and does not sever actin.

The rapid changes in cytoskeletal organization that are mediated by these actin-binding proteins must be achieved without additional protein synthesis, but longer term effects may require additional components or altered levels of existing ones. The level of gelsolin varies during cytoskeletal differentiation. A 50-fold increase in gelsolin synthesis occurs over 48 hours during flattening and spreading of murine cells in response to steroid⁷. Levels of gelsolin messenger RNA also increase about tenfold during differentiation of embryonal carcinoma cells from being small and indistinct to large and flat. By contrast, when early erythroid progenitor cells mature to erythrocytes with a



Schematic representation of different activities of actin and its various actin-binding proteins, together with examples described here.

1. Nadgauda, R.S., Parasharami, V.A. & Mascarenhas, A.F. *Nature* **344**, 335–336 (1990).
2. Anon, *The Sciences* **6**, 9–12 (1967).
3. McKeand, S.E. *J. Am. hort. Sci.* **110**, 619–623 (1985).
4. Smith, D.L., Kelly, K. & Krikorian, A.D. *Physiol. Plant.* **76**, 466–473 (1989).
5. Ammar, S., Benbadis, A. & Tripathi, B.K. *Can. J. Bot.* **65**, 137–142 (1987).
6. Janzen, D.H. *A. Rev. Ecol. Syst.* **7**, 347–391 (1976).
7. Burkill, I.H. *A Dictionary of the Economic Products of the Malay Peninsula* (Ministry of Agriculture and Co-operatives, Kuala Lumpur, 1966).
8. Arber, A. *The Gramineae* (Cambridge University Press, 1934).
9. Ghorai, A. & Sharma, A. *Fed. Rep.* **91**, 281–299 (1980).

highly stable cytoskeleton, the level of this protein is downregulated about 1,000 times in an actin-independent manner⁸. Gelsolin appears to play an important role in these major cytoskeletal changes, but is it unique?

Actin is promiscuous in cells: it can interact with several partners at once. The choice depends on the relative binding affinities and cellular concentrations of the different proteins and on regulatory factors. Because phenotypic changes may depend on the balance of different actin-binding proteins, it is important to assess changes in several of these proteins simultaneously. The levels of gelsolin, profilin, filamin and actin itself have been measured during differentiation of two myeloid cell lines as they acquire the phagocytic and locomotory properties of macrophages⁹. Gelsolin mRNA levels were found to increase 5–10-fold over two days with a corresponding 40–50-fold increase in protein levels, whereas filamin mRNA increased only 2–4-fold with a corresponding 3–7-fold increase in protein levels and there was a greater than two-fold increase in mRNA or protein levels of profilin or actin. The levels of gelsolin, filamin and actin in the differentiated cell lines were similar to those of rabbit macrophages. By contrast, a leukaemic cell line, which differentiates only slightly, showed insignificant changes in the levels of any of these proteins.

These observations emphasize the importance of gelsolin in dynamic rearrangements of the actin cytoskeleton during the spreading and subsequent locomotion of cells, but changes in the levels of other actin-binding proteins may be no less important. Melanoma cell lines lacking filamin show apparently uncontrolled surface blebbing activity together with loss of locomotive capacity¹⁰. Like a car with the clutch disengaged, the engine races but no movement results.

Although rearrangements of cytoskeletal architecture can be correlated with the expression of various actin-binding proteins, this does not prove a causative role. Functional requirements have been explored by disrupting the expression of individual actin-binding proteins. This approach has been particularly successful for myosin using *Dictyostelium discoideum*, an amoeboid cell showing well-defined locomotory behaviour and a developmental cycle. Clear evidence has emerged for the importance of myosin II (analogous to skeletal muscle myosin) in cytokinesis¹¹ and the capping of surface receptors in these cells¹², but cells devoid of this myosin are still capable of cytoplasmic organelle movement, chemotaxis and locomotion. The single-headed globular myosin, myosin I, which does not form filaments and is concentrated at the leading edges of amoebae in motion, may be responsible for these motile activities¹³. ►

Sticklebacks on red alert

WHAT is the advantage to a female of being choosy about her mating partner? Of the several possible answers to this question, one in particular receives support from the work of M. Milinski and T. C. M. Bakker described on page 330 of this issue.

One explanation of sexual selection by female choice is that the male traits preferred are arbitrary, in the sense that they

the results of a study demonstrating that, in the three-spined stickleback *Gasterosteus aculeatus*, females prefer males with the brightest red on their bellies (a trait that is very apparent in the males of the species when in breeding condition, as the picture shows) and that this preference may have evolved as a way of avoiding parasitized males. It seems, perhaps not surprisingly,

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Love-nest — a male three-spined stickleback *Gasterosteus aculeatus* inducing a female to spawn in the nest he has built for the purpose.

confer no direct fitness benefit on offspring, and that they have evolved along with the female mating preference by a kind of 'runaway process'. According to this idea, first suggested by the geneticist R. A. Fisher, the male offspring of preferential matings tend to inherit their fathers' attractive traits, in turn making them more attractive to females when they reach maturity. Female offspring, on the other hand, would tend to inherit their mothers' preferences, the two processes acting together to build up a genetic correlation between preference and preferred trait.

Alternatively, female preferences for particular traits may have come about because of nonadaptive (or even maladaptive) correlates of the otherwise adaptive wiring of the brain. For an example where this may have occurred, see M. J. Ryan *et al.*'s study of sexual selection by 'sensory exploitation' in the tungara frog (*Nature* 343, 66–67; 1990), the conclusion of which is that sensory bias unrelated to any 'runaway process' may be responsible for the female preference to particular male courtship sounds, or 'chucks'.

It is a third explanation — that the traits preferred by females serve to advertise the superior 'quality' of the males, where the 'quality' may be a heritable fitness component — that the report by Milinski and Bakker bears upon. The authors describe

that in this species the intensity of the red breeding coloration correlates with the physical condition of the males, and that males infected by the ciliate parasite that causes 'white spot' suffer a decrease in the intensity of their coloration leading to a reduction in their attractiveness to females. The female preference for brighter, redder males apparent in white light breaks down in green light; together with the results of other experiments, this observation leads the authors to the conclusion that it is the red coloration, and not the presence or absence of parasites as such, that is the cue used when selecting mates.

As Milinski and Bakker readily admit, their study ignores many important aspects of courtship and mating in the three-spined stickleback and is not the last word on female choice in this species. It is not yet known whether in the wild males with brighter coloration tend to be those that are less parasitized or whether coloration is a reflection of heritable differences between males. Furthermore, in this species males care for the eggs and fry until several days after spawning. Therefore, females may avoid parasitized males not because of any genetic benefit that might accrue to their offspring, but because such males may be less capable of paternal care than healthier males.

Rory Howlett

Unfortunately, similar experiments with other actin-binding proteins have not yielded such clear-cut results. A strain of *Dictyostelium* that failed to produce severin showed wild-type motility and developmental behaviour, as did mutant strains defective in α -actinin or gelation factor¹⁴. Although these proteins do not seem to be essential, locomotion and chemotaxis are such vital functions that overlapping activities may have evolved to ensure survival. Bray and Vasiliev have argued in News and Views¹⁵ that actin-based motility has properties similar to parallel processing in computing, in that both depend on the combined activities of the system as a whole. According to this view, it would be necessary to disrupt several actin-binding proteins simultaneously to observe significant effects; any subtle changes caused by the loss of one component may be beyond current resolution. Alternatively, the key actin-binding proteins involved in locomotion have yet to be identified.

Another system that promises to give insights into cytoskeletal function, and that can also be genetically manipulated, is yeast. Yeast cells have an actin distribution that changes during development, but unlike *Dictyostelium* they do not crawl. Amatruda *et al.*¹ in this issue report their analysis of cells in which the gene (CAP2) encoding for the β -subunit of the actin-capping protein (related to *Cap 32/34* of *Dictyostelium*) has been disrupted by insertion or deletion. In contrast to experiments in which severin was ablated in *Dictyostelium*, loss of this protein results in abnormal distribution of actin and therefore in phenotypic changes; the cells grow more slowly and can be three or four times larger. The distribution of chitin in dividing cells, which is normally restricted to the mother-bud junction, is no longer localized but is spread over the entire cell surface, suggesting that CAP2 cannot be replaced by any other protein. Do the differences between these experiments and those in *Dictyostelium* reflect a simpler cytoskeletal system in yeast, which lacks redundancy? Interestingly, mutants lacking profilin¹⁶ also show altered actin distribution and cell morphology, reinforcing this idea. This may not be unexpected because the cytoskeleton in yeast is principally concerned with shape and internal cell movements, whereas in *Dictyostelium* and other motile cells it has a more complex role.

How can these different results be reconciled? The straightforward interpretation of the observations that deletion of filamin in melanoma cells produces dramatic effects, whereas ablation of gelation factor in *Dictyostelium* has no effect, is that these two proteins have totally different cellular functions, despite the similarities between their actin-binding domains. We believe it is more likely

that these proteins share common functions and that the variations seen in these experiments reflect differences in the construction of the cytoskeletal apparatus, how it is used and the extent of redundancy in its regulation. Given these problems, it may be better to study proteins whose distribution is restricted to a particular cell type.

Villin, a component of brush-border cells, is essential in the formation of microfilament bundles and is the first major actin-binding protein to appear at the site of microvilli assembly. During differentiation of the human colon carcinoma cell line HT29 into a well ordered brush-border epithelium, villin mRNA increases about ten times to levels similar to those in human intestinal mucosa¹⁷. In their recent experiments, Friederich *et al.*² transfected fibroblasts (which do not normally contain villin) with villin complementary DNA, and found that the distribution of villin varies with cell density. At low levels of expression in isolated cells, villin was mainly in ruffling membranes at the leading edge, whereas in confluent cells it occurs in microvilli on the dorsal surface of the cells, with no alteration of the actin cytoskeleton on the ventral surface. At high levels of synthesis the effect is much more dramatic. Stress fibres vanish and a profusion of long microvilli appear on the dorsal surfaces of the cells.

Given the limitations of our understanding of motile processes occurring in cells, studies of actin-binding proteins *in vitro* will remain invaluable, particularly because the three-dimensional structure of G-actin at high resolution has been solved¹⁸. As more actin-binding proteins

are cloned and expressed^{19,20}, the nature of their actin-binding sites will be investigated by site-directed mutagenesis and in competition experiments. Structural and functional relationships between different parts of different proteins will be analysed by the construction of chimaeras. But the ultimate goal of deciphering cytoskeletal function can be achieved only by studying cells. □

Michael Way and Alan Weeds are in the MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

1. Amatruda, J.F., Cannon, J.F., Tatchell, K., Hug, C. & Cooper, J.A. *Nature* **344**, 352–354 (1990).
2. Friederich, E., Huet, C., Arpin, M. & Louvard, D. *Cell* **59**, 461–475 (1989).
3. Paulin, D., Perreau, J., Jakob, H., Jacob, F. & Yaniv, M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1891–1895 (1979).
4. Blanchard, A., Ohanian, V. & Critchley, D. *J. Muscle Res. Cell Mot.* **10**, 280–289 (1989).
5. Vandekerckhove, J. *Current Opinion Cell Biol.* **1**, 15–22 (1989).
6. Hartmann, H., Noegel, A.A., Eckerskorn, C., Rapp, S. & Schleicher, M. *J. biol. Chem.* **264**, 12639–12647 (1989).
7. Dieffenbach, C.W., SenGupta, D.N., Krause, D., Sawzak, D. & Silverman, R.H. *J. biol. Chem.* **264**, 13281–13288 (1989).
8. Hinssen, H., Vandekerckhove, J. & Lazarides, E. *J. Cell Biol.* **105**, 1425–1433 (1987).
9. Kwiatkowski, D.J. *J. biol. Chem.* **263**, 13857–13862 (1988).
10. Cunningham, C., Byers, R., Hartwig, J., Stosel, T. & Kwiatkowski, D.J. *Cell Biol.* **109**, 277a (1989).
11. Manstein, D.J., Titus, M.A., De Lozanne, A. & Spudich, J.A. *EMBO J.* **8**, 923–932 (1989).
12. Paternak, C., Spudich, J.A. & Elson, E.L. *Nature* **341**, 549–551 (1989).
13. Fukui, Y., Lynch, T.J., Brzeska, H. & Korn, E.D. *Nature* **341**, 328–331 (1989).
14. Gerisch, G., Segall, J.E. & Wallraff, E. *Cell Motil. Cytoskel.* **14**, 75–79 (1989).
15. Bray, D. & Vasiliev, J. *Nature* **338**, 203–204 (1989).
16. Haarer, B.K. *et al. J. Cell Biol.* **110**, 105–114 (1990).
17. Pringault, E., Arpin, M., Garcia, A., Finidori, J. & Louvard, D. *EMBO J.* **3**, 119–124 (1986).
18. Kabasch, W. & Holmes, K.C. *Biophys. J.* **57**, 397a (1990).
19. Kwiatkowski, D.J., Janmey, P.A. & Yin, H.L. *J. Cell Biol.* **108**, 1717–1726 (1989).
20. Way, M., Gooch, J., Pope, B. & Weeds, A.G. *J. Cell Biol.* **109**, 593–605 (1989).

OZONE DEPLETION

Bad news from the Arctic with signs of dry nitrogen removal

R. L. Jones

POLAR stratospheric clouds play two crucial roles in polar ozone depletion. First, chemical reactions on the surfaces of stratospheric particles result in the liberation of reactive chlorine compounds from their relatively inert reservoir forms which are then available to destroy stratospheric ozone efficiently. Second, gravitational settling of cloud particles containing nitric acid removes nitrogen compounds from the stratosphere — these compounds could otherwise neutralize the available reactive chlorine and slow the rate of any ozone destruction. On page 321 of this issue¹, Fahey *et al.* describe observations showing that our understanding of this second process, and therefore our capacity to predict global ozone loss, is far from complete.

In an earlier report Fahey and his colleagues obtained measurements of H₂O, N₂O and total reactive nitrogen (NO_x) over Antarctica during the 1987 Airborne Antarctic Ozone experiment². These observations showed that up to 80 per cent of the NO_x expected to be present within the Antarctic polar vortex was absent; the vortex was, as it is termed, denitrified. The same air masses were also found to be partially dehydrated, with about half of the water vapour removed. The observation of denitrification was telling in that, in an environment poor in nitrogen oxides, high concentrations of chlorine monoxide — and thus high rates of ozone loss — can be sustained for much longer. So denitrification is an important factor in the development of the ozone

hole over the Antarctic.

Although the general idea of particle sedimentation as a mechanism for such denitrification was accepted, the details were not understood. The occurrence of dehydration and denitrification simultaneously, however, was considered important.

In the stratosphere, water vapour is about 300 times more abundant than nitric acid (1.5–6 parts per million by volume of water vapour as opposed to 5–15 parts per 10^9 by volume of nitric acid). As temperatures fall in the low polar stratosphere (15–25 km high) during the winter months, a temperature threshold (several degrees above that at which ice forms — the frost point, at around 190 K) is reached at which condensation begins. The particles formed at this point are thought to be made of nitric acid trihydrate (NAT), consisting of water vapour and nitric acid molecules in a 3:1 mole ratio^{2,3}. As temperatures fall further, most of the HNO_3 vapour condenses in to the NAT particles, but only a small proportion of the much more plentiful H_2O vapour (a few per cent) is required to co-condense to form NAT. The condensed mass, determined by the available HNO_3 , was thought sufficient to form only small NAT particles (less than 1.0 μm in radius). Such particles would have gravitational fall speeds of the order of metres per day; so at these temperatures, removal of reactive nitrogen would be exceedingly slow.

As temperatures fall below the frost point, however, H_2O vapour condenses. The greater quantity of available H_2O enables much larger ice particles (or the order of 10 μm in radius) to form, with settling speeds of the order of kilometres per day. If HNO_3 could be incorporated into these ice particles, it would then be readily removed from the stratosphere. Thus, it was thought that denitrification and dehydration had to be a single process requiring atmospheric temperatures to fall below the frost point.

This conclusion had serious implications for the question of whether ozone depletion comparable to that in the Antarctic could ever occur in the Northern Hemisphere. The northern winter polar vortex is more disturbed dynamically and warmer than its Antarctic counterpart⁴. As a result, not only are fewer stratospheric clouds expected (and indeed observed⁵) to form in the Arctic, but because temperatures below the frost point were much less widespread and less frequent, it was thought that denitrification would be much less marked in the Northern Hemisphere. With fewer stratospheric clouds, the release of reactive chlorine was expected to be less efficient than over Antarctica, leading to generally lower concentrations of chlorine monoxide. With, in addition, little or no denitrification, the general view was that the return of reactive chlorine to reservoir forms

would proceed correspondingly more quickly in the Arctic. If correct, both arguments would mean that ozone loss on the scale of the Antarctic ozone hole would be highly unlikely in the Northern Hemisphere.

The new results of Fahey *et al.* demonstrate that the above argument is wrong in an important respect. Denitrification and dehydration of the low stratosphere are two separate processes. The observation of denitrification without concurrent dehydration implies that denitrification can occur at temperatures above the frost point, although the authors point out, as have we⁶, that exposure to or proximity to temperatures close to the frost point may be necessary.

So it seems that denitrification in the Northern Hemisphere is spatially more widespread than previously expected on the basis of the Antarctic observations. During the 1988–89 Airborne Arctic Stratospheric Expedition it was also

ORGANELLE INHERITANCE

Redwoods break the rules

Richard G. Harrison and Jeff J. Doyle

PLASTIDS and mitochondria contain populations of covalently closed circular DNA molecules that are replicated and inherited independently of nuclear DNA. Transmission of chloroplast DNA (cpDNA) and mitochondrial DNA (mtDNA) to progeny is determined by maternal and paternal inputs at fertilization and by mechanisms of elimination from the zygote. In addition, because there are many organelles per cell and DNA molecules per organelle, inheritance patterns are also influenced by vegetative sorting at cell division and by differences in replication rates among molecules in germ-cell lineages. A decade ago, research on organelle transmission was limited to the few organisms for which genetic markers were available, and it was possible to postulate 'laws' of organelle genetics¹. Studies of vegetative sorting and differential replication were likewise restricted to a few taxa, because of the paucity of examples of heteroplasmy (single individuals carrying more than one class of mtDNA or cpDNA). The development of restriction-fragment analysis of DNA sequences and the expanding interest of systematic and evolutionary biologists in using organelle DNA molecules as genetic markers has increased enormously the number of taxa for which data on organelle transmission genetics are available. Perhaps predictably, studies of variation in natural populations and comparisons of parents and offspring from controlled crosses have revealed exceptions to most of the earlier generalizations. The latest surprise, reported by

apparent that, despite the less abundant stratospheric clouds, reactive chlorine had been liberated from its reservoir forms over a large fraction of the northern polar vortex⁷. So the potential for ozone loss in northern polar regions is, at least locally, much greater than thought on the basis of the Antarctic data. The limitations in our understanding of the denitrification mechanism highlighted by Fahey *et al.* will have to be overcome if ozone loss in either polar region is to be predicted with any confidence. □

R. L. Jones is at the Meteorological Office, London Road, Bracknell RG12 2SZ, UK.

1. Fahey, D.W. *et al.* *Nature* **344**, 321–324 (1990).
2. Fahey, D.W. *et al.* *J. Geophys. Res.* **94**, 11299–11316 (1989).
3. Hanson, D.R. & Mauersberger, K. *Geophys. Res. Lett.* **15**, 855–858 (1988).
4. Hirota, I., Hirooka, T. & Shiotani, M. *Q. J. R. Met. Soc.* **109**, 443 (1983).
5. McCormick, M.P. *et al.* *J. Atmos. Sci.* **39**, 1387, (1982).
6. Jones, R.L. *et al.* *Geophys. Res. Lett.* (in the press).
7. Brune, W.H. *et al.* *Geophys. Res. Lett.* (in the press).

Neale, Marshall and Sederoff in *Proceedings of the National Academy of Sciences*², is the demonstration of paternal inheritance of mtDNA in redwood trees (*Sequoia sempervirens*).

Using radioactively labelled fragments of cpDNA from *Petunia* and mtDNA from maize, Neale *et al.*² analysed cpDNA and mtDNA restriction fragment length polymorphisms (RFLPs) in parents and offspring from three crosses of coast redwoods. These crosses showed both mtDNA and cpDNA to be inherited paternally. The chloroplast result was not unexpected, because there have been several recent reports of paternal inheritance of cpDNA in conifers^{3–6}. This contrasts with the situation in angiosperms, where maternal inheritance predominates. In neither group of plants, however, is the inheritance pattern strictly uniparental. In conifers, most offspring have only paternal cpDNA, but a few have maternal or novel genotypes^{3,4}. In angiosperms, cytological studies reveal plastid DNA in the sperm or generative cells of 14 per cent of the 200 species sampled⁷, and cases of biparental inheritance have long been recognized⁸. As there is evidence that plastid transmission is influenced both by plastome type and nuclear background, variation in inheritance pattern is perhaps to be expected. Such variation has been documented not only in conifers but also in angiosperms^{9,10}. Given the taxonomic distribution of particular inheritance patterns, mechanisms of uniparental transmission are not likely to be the same for all species⁸.

Much less is known about transmission of plant mitochondria, in part because its complex and structurally labile genome has not provided the kind of easily assayed RFLP markers available for the plastome. Like the *Sequoia* report, the few studies available are based on small samples of progeny, but they suggest that maternal inheritance of the mitochondrial genome predominates even in taxa with paternal or biparental transmission of plastids. Before the *Sequoia* report, there were hints that the rules of mitochondrial transmission were no more absolute than those of the plastid, as some paternal transmission of mtDNA had been observed in barley-rye intergeneric crosses¹¹, and biparental inheritance of mitochondrial RNA had been reported in alfalfa¹². Nevertheless, the finding of strict paternal inheritance of mtDNA among all of the dozen or so progeny analysed in two intra-specific crosses of redwood is news indeed. As the authors themselves point out, the available evidence does not rule out some maternal transmission. Much larger sample sizes would be needed to exclude completely this possibility. But for the first time a species has been found in which it is clear that elimination mechanisms must be postulated for maternal, not paternal mitochondria.

The mitochondrial genomes of animals have received far more attention than those of plants. Animal mtDNA has been characterized as having 'strict' maternal inheritance¹³, although, as in plants, data are available for only a limited array of taxa. Furthermore, the possibility of low levels of paternal leakage cannot be excluded, as there have been remarkably few exhaustive studies of mtDNA inheritance. The apparent rarity of heteroplasmy had long been taken as evidence for maternal inheritance, but even now, when reports of heteroplasmy are commonplace¹³, most examples do not suggest biparental transmission. In virtually all cases, coexisting mitochondrial genomes differ by single changes, often in copy number of tandemly repeated sequences. Heteroplasmy can thus be explained by mutations within homogeneous populations of mtDNA molecules and subsequent drift in genotype frequencies within cell lineages.

In contrast, two recent studies report instances of heteroplasmy in which individual organisms carry two mtDNA genotypes that differ substantially in nucleotide sequence (ref. 14; and W. M. Brown, personal communication). This clearly suggests some paternal mitochondrial transmission in taxa otherwise showing maternal inheritance, and leads to the suspicion that organelle transmission 'rules' may be no more absolute for animals than for plants.

Inheritance patterns are also influenced by changes in the proportions of mtDNA

genotypes in germ-cell lineages. Heteroplasmic ancestors will eventually give rise to a group of homoplasmic descendants as a result of random fixation of alternative types. The number of generations required for fixation appears to be very different in the few organisms for which such data are available¹⁵⁻¹⁷, so it may prove difficult to develop generalizations about intracellular population genetics. For one thing, too few of the parameters, such as effective population size and mutation rate, are known for all but a few taxa. Differential replication will also affect the proportions of two or more organelle types among descendants of a heteroplasmic individual. The success of particular organelle genomes may in part be intrinsic, for example, smaller mtDNA molecules may win the race to replicate¹⁶. But the close cooperation necessary between organelle and nuclear genomes virtually guarantees that such explanations are likely to be at best only part of a more complex story.

The transmission genetics of organelle DNA has become of critical interest to systematists and evolutionary biologists who wish to trace gene genealogies to gain insight into organism phylogenies. Organelle genomes were originally attractive because of their apparent simplicity, uniparental inheritance and absence of recombination, providing 'non-anastomosing maternal lineages' for the interpretation of patterns of variation in natural populations. The *Sequoia* evidence joins the increasing body of exceptions to the rule of maternal inheritance. Although natural instances of organelle genome recombination remain unproven, it will surely not be too much of a surprise to find that organelle lineages anastomose as well. □

Richard G. Harrison and Jeff J. Doyle are in the Division of Biological Sciences, Cornell University, Ithaca, New York 14853, USA.

1. Birky, C.W. Jr *Am. Rev. Genet.* **12**, 471-512 (1978).
2. Neale, D.B., Marshall, K.A. & Sederoff, R.R. *Proc. natn. Acad. Sci. U.S.A.* **86**, 9347-9349 (1989).
3. Neale, D.B., Wheeler, N.C. & Allard, R.W. *Can. J. For. Res.* **16**, 1152-1154 (1986).
4. Szmidt, A.E., Alden, T. & Hallgren, J.-E. *Plant molec. Biol.* **9**, 59-64 (1987).
5. Neale, D.B. & Sederhoff, R.R. *Theor. appl. Genet.* **77**, 212-216 (1989).
6. Stine, M., Sears, B.B. & Keathley, D.E. *Theor. appl. Genet.* **78**, 768-774 (1989).
7. Corriveau, J.L. & Coleman, A.W. *Am. J. Bot.* **75**, 1443-1458 (1988).
8. Sears, B.B. *Plasmid* **4**, 233-255 (1980).
9. Chiu, W.-L., Stubbe, W. & Sears, B.B. *Curr. Genet.* **13**, 181-189 (1987).
10. Schumann, C.M. & Hancock, J.F. *Theor. appl. Genet.* **78**, 863-866 (1989).
11. Soliman, K., Fedak, G. & Allard, R.W. *Genome* **29**, 867-872 (1987).
12. Fairbanks, D.J., Smith, S.E. & Brown, J.K. *Theor. appl. Genet.* **76**, 619-622 (1988).
13. Harrison, R.G. *Trends Ecol. Evol.* **4**, 6-11 (1989).
14. Satta, Y. et al. *Genet. Res.* **52**, 1-6 (1988).
15. Solignac, M., Genereux, J., Monnerot, M. & Mounolou, J.-C. *Molec. Gen. Genet.* **197**, 183-188 (1984).
16. Rand, D.M. & Harrison, R.G. *Genetics* **114**, 955-970 (1986).
17. Ashley, M., Laipis, P.J. & Hauswirth, W.W. *Nucleic Acids Res.* **17**, 7325-7331 (1989).

Steering the snow

Snow, says Daedalus, is a mixed blessing. A snowy landscape is lovely to look at, and highly profitable for the ski-ing industry, but snow is a major nuisance on sports-grounds, racecourses, runways, and roads. It has to be melted away with massive dressings of salt — which does damage of its own, by crumbling concrete and rusting vehicles. Now Daedalus plans to tame this unruly form of winter weather.

Any brute-force attempt to melt or hold back or pre-precipitate a snowstorm is, of course, doomed to failure; atmospheric energies are far too great. Daedalus's 'Snow-steerer' accepts whatever snow nature provides, and merely steers it in the most convenient direction.

It exploits the phenomenon of the Rayleigh disc, that neat device which detects a sound wave by turning at right-angles to its direction of propagation. Such a disc responds to a very small sound energy, and absorbs only a tiny fraction of that. Most snowflakes have a vaguely disc-like structure, derived from the flat hexagonal crystal habit of ice itself, and should turn at right-angles to a suitable sound beam. Furthermore, a disc tends to fall obliquely, gliding unstably sideways parallel to its diameter. So a beam of sound briefly pulsed from the Snow-steerer should turn all the snowflakes in its vicinity parallel to its diverging wave-front. They will then glide slantwise out of the beam, away from the Snow-steerer.

The effect will be sadly transient. The asymmetry of the snowflakes and the instability of their glide will soon randomize their fall again. To stabilize their orientation and keep them on their glide-path out of its sound beam, the Snow-steerer will have to sound repeatedly or even continuously. Mercifully, snowflakes are small enough to be orientated by inaudibly ultrasonic wavelengths, so even a very loud snow-steering sound would not annoy a human population. (Bats might be upset, but they have no business to be flying in a snowstorm anyway.)

The Snow-steerer will revolutionize our battles with winter weather. Even in the fiercest snowstorm, the ground around it will remain effortlessly clear. Set along the central barriers of motorways, Snow-steerers will keep the traffic safely flowing. Installed in the middle of a football field, they will repel the snow away from the pitch and onto the spectators. Conversely, set around a ski-slope, they will concentrate all the local snowfall onto the slope. Desperate slope-owners cursing the greenhouse effect will be able to make the best use of their fast-wasting assets. And by subtle control or programming of a Snow-steering array, an ingenious sculptor might create the world's first self-accumulating snowman.

David Jones

Profiting from reburial

SIR—I wish to comment on Paul Bahn's account in *News and Views*¹ of the First International Congress on the Reburial of Human Remains. The meeting was a scientific sham, a media event for the advocates of reburial². Groups supporting preservation were excluded. Archaeologists and other researchers (including the forensic specialist Clyde Snow), who were scheduled to speak, were barred from the podium by the activists who took control of the meeting.

The reburial movement in the United States, often motivated by media attention, politics and money, masquerades as a religious or racial rights movement. But it has little concern for individuals' religious freedom or the rights of the true next-of-kin to claim their ancestors' remains³. At stake are large quantities of money being paid to 'native' site watchers and advisors (the \$1,000,000 recently paid to 'Indian watchers' for the Santa Barbara oil pipeline project, for example⁴).

Public money for reburials is the latest growth industry for numerous activists; \$135,000 of taxpayers' money was used to pay off land-owners, lawyers, archaeologists and activists in an effort to bury 146 poorly studied skeletons in Salina, Kansas⁵. Religious and historic traditions, accurate identifications and the desires of the next-of-kin have little influence on its many activists who demand reburial for all remains under a variety of self-styled 'traditional' religions⁶. Thus Stanford University has released 550 Ohlone skeletons to individuals who had identified themselves with this tribe (the last recognized member of which died in the early 1800s⁶).

Political compromises and legislation by misguided politicians, are bringing archaeology to a halt and are destroying many crucial collections of skeletons and artefacts. Yet many government scientists are silenced by intimidation, while museums are being forced to dispose of their collections. The Smithsonian Institution expects to lose a quarter of its skeleton collection in exchange for new building and more federal funding⁷. Yet professional scientific groups have mounted, for the most part, a poor and ineffective defence.

Reburial, with few exceptions, robs the world of important records of environment, life style and human endeavour. It seldom reflects the wishes of the direct descendants. These losses to science, as well as of money which, otherwise, could be used for better health care and education, will seriously affect medical research into cancer, diabetes and arthritis (which affect Native Americans more than any other faction of the general public) as well as cultural studies⁸. It will also destroy the

heritage of many prehistoric peoples, retard land-claim litigation and narrow religious rights and freedoms.

This loss to the world, caused by greed, ignorance and shortsighted zealotry, will unfortunately harm the very people it is meant to help. Custody of remains should be given to only the biologically determined, direct next-of-kin. Others should be kept in perpetuity for study.

E.J. NEIBURGER

*Ethnic Minority Council of America,
33263 North Cove Road,
Grayslake,
Illinois 60030,
USA*

1. Bahn, P. *Nature* **342**, 123–124 (1989).
2. Gregg, J. *Paleopath. Newsletter* **67**, 7 (1989).
3. *Argus Leader* (Sioux Falls, SD) p. 6E (7 August, 1988).
4. *Santa Barbara News* p. A8, B3. (3 March, 1987).
5. *Kansas City Star* (26 February, 1989).
6. *Time* p. 29 (10 July, 1989).
7. Inouye, D. *US Senate Bill* 978 (May 1989).
8. *Trends in Indian Health 1989* (US Indian Health Service, 1989).

Looking for a link

SIR—Aldhous's *News report*¹ does much to shed light on the controversial issue of the relationship between bovine spongiform encephalopathy (BSE) and its human counterpart, Creutzfeldt–Jakob disease. BSE is an impressive recent example of interspecies transfer of an infectious agent, in this case the apparent transfer of scrapie from sheep to a new host species, cattle.

Remarkably, a similar transfer seems to have occurred before in at least one other species under much the same circumstances, resulting in the disease now termed transmissible mink encephalopathy (TME). TME was first identified in 1947 on a mink ranch in Wisconsin², and further outbreaks were reported in 1961 and 1963. Like BSE, these outbreaks resulted in high fatality rates among adult animals on affected farms. Investigation of the outbreaks³, supported by additional evidence from a 1965 outbreak in Finland⁴, suggested a probable origin from feed which incorporated sheep by-products and was putatively contaminated with scrapie.

As beef by-products were also included in the feed, transmission via infected cattle, rather than directly from sheep, cannot be excluded. Experimental inoculation of mink with the agent causing scrapie resulted in a TME-like spongiform encephalopathy⁴, confirming an association between scrapie and TME, although not all strains of scrapie caused disease in mink^{3,5}. Scattered outbreaks of TME have been identified since the 1960s in Canada and Europe⁵, but TME has not recently been reported in the United States.

Although interspecies transfer of infec-

tious agents is believed to be of great importance in the emergence of disease⁶, most are recognised long after the fact, and so little is known about specific events and mechanisms involved in transfer. BSE, and the earlier transfer of TME, are among the few well-documented instances. Others may be the morbillivirus implicated in seal deaths in the North Sea and canine parvovirus 2. As Aldhous points out¹, Creutzfeldt–Jakob disease is rare even though scrapie has been known in sheep for at least two centuries, and the possibilities of human infection in a manner comparable to the acquisition of BSE by cattle have presumably been high.

The factors causing outbreaks of diseases like BSE and TME in cattle and mink are unknown, nor is it clear why such outbreaks have apparently not occurred in humans despite probable exposure to the same infectious sources. Further research on BSE may help to provide some of the answers.

STEPHEN S. MORSE

*The Rockefeller University,
1230 York Avenue,
New York, New York 10021-6399, USA*

1. Aldhous, P. *Nature* **343**, 196 (1990).
2. Hartsough, G.R. & Burger, D. *J. infect. Dis.* **115**, 387–392 (1965).
3. Marsh, R.F. & Hanson, R.P. in *Slow Transmissible Diseases of the Nervous System* (eds Prusiner, S.B. & Hadlow, W.J.) Vol. 1, 451–460 (Academic, New York, 1979).
4. Hanson, R.P. *et al. Science* **172**, 859–861 (1971).
5. Marsh, R.F. in *Slow Virus Diseases of Animals and Man* (ed. Kimberlin, R.H.) 359–380 (Elsevier/North-Holland, Amsterdam, 1976).
6. Morse, S.S. & Schluederberg, A. *J. infect. Dis.* (in the press).

Origins of segment periodicity

SIR—The analysis of *Hox* genes in vertebrates has revived the old problem of the origin of segmentation. The consensus¹ is that "segmentation seems to have arisen separately in the lineages that lead to the arthropods and vertebrates". While this is the position taken by most textbooks (and assumed to be correct by most readers), it is clear that the lines of evidence on which this view is based are all quite arguable. A second, more serious problem is that the insight provided by the genetic analysis of segmentation in *Drosophila* is not, as yet, to be found in zoology textbooks.

A chief conclusion of this genetic analysis is that the primary event in segmentation is the establishment of a periodic expression of certain genes along the anteroposterior axis (see ref. 2). The first detectable signs of ectodermal segmentation occur long after this periodic organization has taken place. Furthermore, the manifestation of periodicity may differ according to tissue or developmental stage: for example, the domains of action of homeotic genes are parasegmental, that is, out of phase with the segments them-

selves, both in the embryonic epidermis and in the adult central nervous system (CNS).

These results suggest that the earliest signs of morphological segmentation (for example, somite formation, in the case of vertebrates) may be very poor indicators of the basic mechanisms which generate periodicity, and that it would be unsafe to rely on such criteria to assess the phylogenetic origin of segmentation. Indeed, if the importance of periodic organization is recognized, then the question whether segmentation did or did not arise independently in different phyla becomes irrelevant.

An alternative approach to the question would be to look for the earliest signs of periodic organization in metazoans. Architecture of the CNS across the major phyla shows that periodic organization of the CNS is a common feature of most of them, including some of the most primitive flatworms³. Furthermore, periodicity is obvious in the CNS of primitive molluscs (chiton) and in echinoderms (starfish), even though overt body segmentation is lacking or doubtful in these organisms.

This suggests that the principle of periodicity (or metamerization) may have emerged in the CNS before it was applied to any other tissue, possibly to help set up a scaffold of increasingly complex stereotyped connections. Indeed, basic determinants of the CNS structure, such as the paths laid down by pioneer neurons, have evolved remarkably little over most of the arthropod lineage⁴, presumably because of the structural constraints involved in the establishment of a stereotyped pattern of connections.

In this view, periodic organization would definitely have arisen (in the CNS) before the separation of the major lineages. Recent work on the segmentation of the hindbrain in vertebrates⁵ and its correlation to the realms of action of *Hox* genes⁶ are certainly consistent with this view.

ALAIN GHYSEN

Université Libre de Bruxelles,
Laboratoire de Genetique,
Rue des Chevaux 67,
1640 Rhode-Saint-Genèse, Belgium

1. Wilkinson, D.G. et al. *Nature* **341**, 405–409 (1989).
2. Ingham, P. *Nature* **335**, 25–34 (1989).
3. Bullock, T.H. & Horridge, G.A. *Structure and function in the nervous system of invertebrates* (Freeman, San Francisco, 1965).
4. Thomas, J.B. et al. *Nature* **310**, 203–207 (1984).
5. Lumsden, A. & Keynes, R. *Nature* **337**, 424–428 (1989).
6. Murphy, P. et al. *Nature* **341**, 156–159 (1989).

Genetic linkage and psychiatric disease

SIR—The claim by Robertson in her News and Views article¹ that there is no persuasive evidence linking any psychiatric disease to a single locus is incorrect. There is a well-established linkage for manic depression on the X chromosome, which was first reported 20 years ago² and has since been confirmed in Sardinia, Belgium, Israel and the United States in independent samples^{3–7}. In the Israeli sample the lod score was greater than 9.00, which is highly significant. When all the studies showing X linkages are combined, the lod is greater than 15.

Robertson also suggested that the chromosome 5 linkage for schizophrenia reported from our laboratory was incorrect: her comment was based on a misleading interpretation of our unpublished data by Ken Kidd at *Nature's* recent conference on molecular genetics and Human Disease. The marker M4, which lies inbetween the two markers that we have already published as being linked to schizophrenia, does in fact show linkage to schizophrenia with a lod of 1.53. The reason that it is not higher than this is because it is not providing as much information from the seven families as the other two markers.

This is one of the difficulties commonly encountered in linkage research because it cannot be guaranteed that key individuals included in an analysis will be heterozygous for the marker polymorphisms and thus become informative for linkage. Unfortunately this is due to chance. In our families, M4 is occasionally informative in the same individuals as the marker p105–599Ha, which shows a two-point lod to schizophrenia of 4.95 for the broadest definition of affection. Where it is informative, M4 gives exactly the same linkage information as p105–599Ha, except that in one family (F20) two individuals who were previously uninformative now become cases of incomplete penetrance. We originally argued⁸ that the presence of incomplete penetrance for schizophrenia would make it difficult to estimate accurately the recombination fraction between linkage markers and the chromosome 5 schizophrenia locus, unless our 'penetrance-free' model was used.

Robertson also states that readers of *Nature* should be sceptical of linkages to psychiatric disorders that have not reached a lod of 6. There is no reason why

psychiatric disorders should be singled out for special treatment, except for the fact that some psychiatric disorders show incomplete penetrance, and that the penetrance must be estimated jointly with the recombination fraction. If this is done, then the appropriate level of significance is a lod of 3.85. The presence of heterogeneity of linkage, uncertainty about definitions of caseness, and the presence of phenocopies very much reduces the chance of spurious linkage being reported for psychiatric disorders and this should alert us to the possibility that false negative linkages are more likely to occur than spuriously positive ones.

HUGH GURLING

Molecular Psychiatry Laboratory,
University College and Middlesex School
of Medicine,
University of London,
W1P 7PN UK

SIR—Miranda Robertson's lucid discussion¹ of the problems already generated by the application of linkage methods appropriate to mendelian disorders to disorders that rarely show evidence of mendelian segregation overlooked two of the central quicksands on which the foundation of this claim rested, as indeed did your referees of the paper of Egeland et al. in 1987 (ref. 2).

First, in the analysis of a single marker, it is not sufficient to consider only the prior odds against the occupancy of a locus by a particular allele, which in man is of the order of twice the number of chromosomes, so that odds of a thousand-to-one, even if calculated after defining the maximum, is likely to be in error in only about 5 per cent of apparent linkages. It is also necessary to consider the prior odds against such a simple mechanism: the prior expectation that a functional disorder with very variable manifestation in an organ of extreme complexity of structure and function will have a simple mechanism must have some consideration, even if highly subjective. The possibility of two or more strongly predisposing alleles at different loci is at least more likely than one, but as there are more than a thousand ways in which two loci can be chosen with fairly distinct segregation, this approach is also limited.

The Amish study was complicated by a more serious problem, one underlying the optimistic predictions of the scope of linkage in the original article of Botstein et al.³ that spawned the term "the new genetics"⁴: the more marker loci are tested, the more likely the discovery of a false linkage due to fortuitous segregation. The Amish study of 1987 was based on a previous trawl of the genome with twenty markers. When one of these was found to cosegregate, a set of markers known to be closely

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. They need not arise out of anything published in *Nature*. In any case, priority will be given to letters of fewer than 500 words and five references. □

1. Robertson, M. *Nature* **342**, 222 (1989).
2. Winokur, G., Clayton, P.J. & Reich, T. *Manic Depressive Illness* (Mosby, St Louis, 1969).
3. Mendlewicz, J., Fleiss, J.L. & Gieve, R.R. *J. Am. med. Assn* **222**, 1624–1627 (1972).
4. Del Zompo, M. et al. *Acta Psychiat. Scand.* **70**, 828–837 (1984).
5. Risch, N. & Baron, M. *Ann. hum. Genet.* **46**, 153–166 (1982).
6. Baron, M. et al. *Nature* **326**, 289–292 (1987).
7. Mendlewicz, J. et al. *Lancet* **i**, 1230–1231 (1987).
8. Sherrington, R. et al. *Nature* **336**, 164–167 (1988).

linked was used, and this also cosegregated, as indeed it could not fail to do since the same family was analysed and because previously informative meioses were not excluded.

This type of preliminary trawl followed by detailed study has recently been recommended by Lander⁵, who writes:

In the future, we envisage that linkage studies will initially employ a standard battery of perhaps 100–200 highly informative RFLPs distributed throughout the genome. Any interval showing suggestive evidence of linkage will then be studied with a higher density of RFLPs. This will extract the full information from the pedigrees, and will increase the LOD score.

This procedure will certainly increase the lod score, but it would be more accurate to state that it will also extract full mis-information, and there is no way in which information and mis-information can be separated unless different families are used.

Lander summarizes the problem lucidly by stating that "if the trait is, in fact, more complicated than assumed, linkage will not be detected. A negative result in a complete genome search will at least prove that the disease is more complex than had been assumed". While amentia and dementia are good grist for the mill now available for the shotgun geneticist, it would be useful to know who the assumers are for simplicity in the psychoses.

J. H. EDWARDS

Genetics Laboratory,
Department of Biochemistry,
University of Oxford,
South Parks Road, Oxford OX1 3QU, UK

1. Robertson, M. *Nature* **342**, 222 (1989).
2. Egeland et al. *Nature* **325**, 783–787 (1989).
3. Botstein, D. et al. *Amer. J. Hum. genet.* **32**, 314–331 (1980).
4. Comings, D.A. *Amer. J. Hum. genet.* **32**, 314 (1980).
5. Lander, E.S. in *Genome Analysis* (ed. K.E. Davies) (IRL, Oxford, 1989).

Analogous alternative splicing

SIR—The two molecular forms of the A chain of platelet-derived growth factor (PDGF), which arise by alternative modes of mRNA splicing, differ in length by a 15-residue C-terminal extension rich in basic amino acids which is derived from a separate sixth exon^{1,2}. The differently spliced mRNAs have so far been found in

chains of PDGF might serve as nuclear targeting sequences if the N-terminal signal sequences for secretion were deleted^{8,9}. But it is difficult to imagine a function for the PDGF chains in the nucleus as well as a normal route by means of which the molecules would escape secretion. Instead, it is possible that the

hPDGF-A	194	G R P R E S	G K K R K R K R L K P T	end 211
xPDGF-A	198	G F F T S P A L V L T G R T R E T	G K K Q K R K K L K P T	end 226
hVEGF/VPF	116	K S V R G K G K	G Q K R K R K K S R Y K S W S V	cont. to 189
conserved basic residues		*	* * * * *	

Amino acid sequence homology in the regions of human PDGF-A, *Xenopus* PDGF-A and human VEGF/VPF encoded by the sixth exon. The longer versions of both human and *Xenopus* PDGF end with this region while the molecule of VEGF/VPF continues for a further 50 residues.

all cell lines and tissues expressing PDGF (ref. 4), although the shorter form invariably predominates. The recognition that two forms persist even in the amphibian *Xenopus laevis* shows that the process of alternative splicing, involving an exon corresponding to human exon 6, is well-conserved in evolution⁵.

Recent reports of the cloning of an endothelial cell mitogen called vascular endothelial growth factor (VEGF)⁶ or vascular permeability factor (VPF)⁷, whose structure is distantly related to those of the A and B chains of PDGF, make it clear that alternative modes of splicing, involving a sequence similar to that of the sixth exon of human PDGF-A, also takes place with this mRNA, and at an analogous position. The figure shows an alignment of the amino acid sequences of the various growth factors encoded by the differently spliced mRNAs. Interestingly, the degree of sequence identity between human PDGF and VEGF/VPF is greater in this region than in any other part of the molecules.

Although the biological function of the basic motif encoded by the sixth exon of the genes for the two molecules is not known, the basic domains of the A and B

basic motif common to the longer variants of PDGF-A and VEGF/VPF is an address for an extracellular target, perhaps one involved in regulating the accessibility of the polypeptides or in directing it to a specific interstitial structure.

CHRISTER BETSHOLTZ
FREDRIK RORSMAN
BENGT WESTERMARK

Department of Pathology,
University of Uppsala,
University Hospital, S-75185 Uppsala,
Sweden

ARNE ÖSTMAN
CARL-HENDRIK HELDIN

Ludwig Institute for Cancer Research,
Biomedical Center,
Box 595, S-75123 Uppsala,
Sweden

1. Betsholtz, C. et al. *Nature* **320**, 695–699 (1986).
2. Rorsman, F. et al. *Molec. cell. Biol.* **8**, 571–577 (1988).
3. Bonthon, D.T., Morton, C.C., Orkin, S.H. & Collins, T. *Proc. nat. Acad. Sci. U.S.A.* **85**, 1492–1496 (1988).
4. Matoskova, B., Rorsman, F., Svensson, V. & Betsholtz, C. *Molec. cell. Biol.* **9**, 3148–3150 (1989).
5. Mercola, M., Melton, D.A. & Stiles, C.D. *Science* **241**, 1223–1225 (1988).
6. Leung, D.W., Cachianes, G., Kuang, W.-J., Goeddel, D.V. & Ferrara, N. *Science* **246**, 1306–1312 (1989).
7. Keck, P.J. et al. *Science* **246**, 1309–1312 (1989).
8. Lee, B.A., Maher, D.W., Hannink, M. & Donoghue, D.J. *Molec. cell. Biol.* **7**, 3527–3537 (1987).
9. Maher, D.W., Lee, B.A. & Donoghue, D.J. *Molec. cell. Biol.* **9**, 2251–2253 (1989).

Origins of T-cell leukaemia virus

SIR—The causal retrovirus of adult T-cell leukaemia (ATL), termed HTLV-1, is also responsible for HTLV-1-associated myelopathy (HAM) and tropical spastic paraparesis (TSP). So much is clear from investigations in Martinique, Colombia and Jamaica and, more recently, in Japan. But the origin of the carriers of HTLV-1, and the route by which they reached South America, remain obscure.

Gallo's suggestion¹ that the origin of HTLV-1 was in Africa, and that it reached Japan and the Americas by the slave trade, has been vigorously disputed^{2–4}. According to Hinuma, one of the co-discoverers of the virus, HTLV-1 may first have infected human beings some thousands of years ago, when the infection would not have been fatal because people did not live long enough to develop ATL⁵.

It is significant that there is a high frequency of HTLV-1 carriers among the Ainu people of Hokkaido, who are commonly regarded as the descendants of the native population living mostly in northern Japan in the pre-agricultural Jomon period more than 2,500 years ago⁶. Evidence of a high frequency of HTLV-1 carriers among the aborigines of Papua New Guinea and in several Melanesian islands suggests a route between Asia and South America, yet the main endemic focus of infection in Colombia is a group of blacks of African origin living on the South Pacific coast, which would support an African origin of the virus.

On the other hand, cases of TSP positive for HTLV-1 have been found in Chile. (5 of 34 patients were of mixed white and araucano descent.) We have also found three South American Indian women of the Páez group sick with TSP and carrying HTLV-1 in an isolated region of the Colombian Andes at more than 2,000 m above sea-level. These, and the similar cases reported from the high valleys of the Peruvian Andes, make it difficult to accept an exclusive African origin for South American HTLV-1. We think it more likely that migration from Asia carried the virus to this continent.

VLADIMIR ZANINOVIC

Universidad del Valle
Cali, Colombia

TOMÁS ZAMORA

Universidad del Cauca,
Popayán, Colombia

KAZUO TAJIMA

Aichi Cancer Center,
1-1 Kanokoden, Chikusa-Ku,
Nagoya 464, Japan

1. Gallo, R.C. *Scient. Am.* **255**, 88–98 (1986).
2. Kantra, S.S. *Nature* **321**, 733 (1986).
3. Ishida, T. *Nature* **322**, 504 (1986).
4. Taguchi, H. *Nature* **323**, 764 (1986).
5. Hinuma, Y. et al. *Int. J. Cancer* **29**, 631 (1982).
6. Ishida, T. et al. *J. Infect.* **11**, 153–157 (1985).

Cosmopolitan underwater fauna

SIR—In their interesting letter published in Scientific Correspondence, Smith *et al.*¹ reported their discovery of a chemosynthetic community on a whale carcass, and suggested that the whale's skeleton provided a source of sulphides to fuel their metabolism. But this speculative idea that this chemoautotrophic fauna is similar to hydrothermal vent faunas is not supported by other observations.

The *Beggiatoa* spp. mats described by Smith *et al.*, for example, are ubiquitous on sulphide-rich environments such as intertidal mudflats, anoxic basins and fish farms^{2,4}, and are not restricted to vents and seeps. Further, of 219 published species of vent fauna, only 11 have been found in other habitats, including carcasses (V.T., unpublished data). The biogeographical patterns of true vent fauna are more likely to be related to seafloor spreading than by a fortuitous mechanism such as via carcasses of large animals.

In any event, the species described by Smith *et al.* are not all found in vents. The described molluscs exploit sulphide availability and are found in many such habitats. A few, like *C. pacifica*, may have the adaptations to enter vents and seeps. The truly endemic vent animals are not found on this carcass. But we do not believe that a few shared features between vent and whale carcass habitats provide evidence for an important dispersal mechanism.

VERENA TUNNICLIFFE

Department of Biology,
University of Victoria,
British Columbia, Canada V8W 2Y2

S. KIM JUNIPER

Department of Oceanography,
University of Quebec,
Rimouski, Canada G5L 3A1

1. Smith, C.R. *et al.* *Nature* **341**, 27–28 (1989).
2. Jorgensen, B.B. *Phil. Trans. R. Soc.* **B298**, 543–561 (1982).
3. Juniper, S.K. & Brinkhurst, R.O. *Mar. Ecol. Prog. Ser.* **33**, 41–50 (1986).
4. Grant, J. & Bathman, U.V. *Science* **236**, 1472–1474 (1987).

How secretion is inhibited

SIR—It has been suggested¹ that glucose inhibition of glucagon secretion is mediated by GABA activation of chloride channels in the pancreatic α_2 -cells, that the GABA is co-secreted with insulin from the pancreatic β -cells and that any circumstances in which insulin secretion is stimulated will lead to suppression of glucagon release. This suggestion is an interesting possibility which, however, can hardly explain inhibition by glucose.

My scepticism originates from the

physiology of glucose homeostasis. Whereas secretion of blood glucose-lowering insulin begins above 4–5 mM of the sugar, the release of blood glucose-raising glucagon is half-maximally inhibited by such concentrations². The postulated glucose-induced co-secretion of GABA with insulin would consequently start in a concentration range where glucagon secretion is only marginally affected by the sugar.

An alternative explanation for glucose inhibition of glucagon release is a sugar-induced lowering of cytoplasmic Ca^{2+} caused by intracellular sequestration and outward transport^{3,4}. Although the glucose dependence of this phenomenon has yet to be elucidated, it is noteworthy that a similar glucose-induced lowering of cytoplasmic Ca^{2+} , which is a component in the action of the sugar on the pancreatic β -cell, is half-maximal at 6 mM glucose⁵.

A final comment concerns the citation of previous work. The original detection of GABA within the pancreatic islets was not made by Okada *et al.*⁶, although the authors probably believed this at the time. But the discovery was actually made four years earlier in two independent laboratories^{7,8}.

ERIK GYLFE

Department of Medical Cell Biology,
University of Uppsala,
Biomedicum Box 571,
S-751 23 Uppsala, Sweden

1. Rorsman, P. *et al.* *Nature* **341**, 233 (1989).
2. Gerich, J.E. *et al.* *A. Rev. Physiol.* **38**, 353 (1976).
3. Johansson, H. *et al.* *Biochem. Biophys. Res. Commun.* **147**, 309 (1987).
4. Johansson, H. *et al.* *Cell Calcium* **10**, 205 (1989).
5. Gylfe, E. *J. Biol. Chem.* **263**, 13750 (1988).
6. Okada, Y. *et al.* *Science* **194**, 620 (1976).
7. Briel, G. *et al.* *Acta physiol. Scand.* **84**, 247 (1972).
8. Panten, U. *et al.* *FEBS Lett.* **20**, 225 (1972).

Diacylglycerol as messenger

SIR—Smith *et al.*¹ recently reported that the secretagogue glucose decreases the voltage threshold for activation of L-type Ca^{2+} channels in insulin-secreting cells. This finding nicely confirms, in normal rat pancreatic β -cells, results obtained with the carbohydrate secretagogue glyceraldehyde on the rat insulinoma cell line RINm5F (ref. 2). But it provides no information on the nature of the metabolic product supposed to stimulate the channels.

It is well established that Ca^{2+} channel function can be modulated by cyclic AMP-dependent phosphorylation³, but carbohydrate stimuli have little or no effect on cyclic AMP levels in insulin-secreting cells⁴. On the other hand, it is known that glucose and glyceraldehyde evoke rapid *de novo* synthesis of diacylglycerol in normal pancreatic β -cells⁵ and insulinoma cells⁶.

Since the cell-permeable diacylglycerol

analogue didecanoylglycerol has been shown to mimic the effect of carbohydrate secretagogues on gating of single L-type Ca^{2+} channels in insulinoma cells⁷, diacylglycerol may fulfil the required messenger role. Diacylglycerol could work by protein kinase C-mediated phosphorylation, or by an as yet unknown mechanism⁸.

O. H. PETERSEN

MRC Secretary Control Research Group,
Physiological Laboratory,
University of Liverpool,
PO Box 147, Liverpool L69 3BX, UK

C. B. WOLLHEIM

Institut de Biochimie Clinique,
Centre Medical Universitaire,
Universite de Geneve,
1211 Geneve 4, Switzerland

1. Smith, P.A., Rorsman, P. & Ashcroft, F.M. *Nature* **342**, 550–553 (1989).
2. Velasco, J.M., Petersen, J.U.H. & Petersen, O.H. *FEBS Lett.* **231**, 366–370 (1988).
3. Reuter, H. *News physiol. Sci.* **2**, 168–171 (1987).
4. Prentki, M. & Matschinsky, F.M. *Physiol. Rev.* **67**, 1185–1248 (1987).
5. Peter-Riesch, B., Fathi, M., Schlegel, W. & Wollheim, C.B. *J. clin. Invest.* **81**, 1154–1161 (1988).
6. Wollheim, C.B. *et al.* *EMBO J.* **7**, 2443–2449 (1988).
7. Velasco, J.M. & Petersen, O.H. *Q. J. exp. Physiol.* **74**, 367–370 (1989).
8. Kolesnick, R.N. *J. Biol. Chem.* **262**, 16759–16762 (1987).

Copper-bottomed earwax

SIR—The unexplained liking of several cats for human ear wax (*Nature* **343**, 220; 1990) is a reminder that pica, a desire for unusual food, can sometimes be traced to dietary deficiency, such as a lack of iron¹. Cerumen (ear wax) has a high copper content², which may relate to its bactericidal and mycotoxic properties³. Copper occurs in most tissues of the animal body, and minute amounts are necessary for the incorporation of iron into haemoglobin⁴. In the cytochrome system, the function of copper is similar to that of iron. Copper requirements vary among species, and are influenced by the intake of other mineral elements such as iron and molybdenum⁵. If there is a copper deficiency, anaemia is a common sequel⁴.

It may be that some domestic animals experience a copper-deficient diet, leading them to seek supplementary sources of this mineral.

PATRICK M.P. O'CONNOR

Department of Radiology,
Albany Medical Center Hospital,
Albany, New York 12208, USA

1. Reynolds, R.D. *et al.* *Ann. int. Med.* **69**, 435–440 (1968).
2. Yassin, A., Mostafa, M.A. & Moawad, M.K. *J. laryngol. Otol.* **80**, 933–937 (1966).
3. Megarry, S. *et al.* *J. laryngol. Otol.* **102**, 671–672 (1988).
4. Jones, C.J. & Hunt, R.D. in *Veterinary Pathology* 1066–1068 (Lea & Febiger, Philadelphia, 1983).
5. Hays, W.W., Swenson, M.J. in *Dukes' Physiology of Domestic Animals* (ed Swenson, M.J.) 395–412 (Comstock, Ithaca, 1977).

The bear and the pussycat

Richard L. Garwin

A Shield in Space? Technology, Politics, and the Strategic Defense Initiative. Sanford Lakoff and Herbert F. York. *University of California Press: 1989. Pp. 409. \$35.*

A FURTHER subtitle, "How the Reagan administration set out to make nuclear weapons 'impotent and obsolete' and succumbed to the fallacy of the last move", sums up this text. York, an important player on the military, technological and diplomatic scene since the early 1950s, has been a respected observer of SDI since it was revealed by US president Ronald Reagan on 23 March 1983. From the perspective of six years' of SDI experience as contrasted with early rhetoric, and buttressed by the frank views of several of the key participants, the authors have provided a solid and informative account of the SDI programme.

Held in deepest secrecy by the president and a few close aides from those in his administration with the experience and competence to advise on feasibility, cost and perspective, the SDI concept was sketched by Reagan as a vision to "render nuclear weapons impotent and obsolete". Within days, SDI emerged as a National Security Decision Directive and only then was subjected to intensive study by a hand-picked team of government and industrial technologists.

The authors rise to the challenge of chronicling and assessing the SDI programme, which (with unchanged name) has progressed from a principled *replacement* for deterrence of nuclear war to a proclaimed *support* for the traditional approach to deterrence by threat of retaliation. No longer a leak-proof defence against nuclear weapons of all kinds, the SDI umbrella has been resilient enough to accommodate the enlargement of mission from a long-term research programme to "protecting the option for early deployment" to a "requirement" by the Joint Chiefs of Staff for a phase 1 SDI system that reportedly could destroy between 11 and 30 per cent of the warheads expected to be launched by the Soviet Union in a first wave against US strategic targets.

SDI plans and projections are interpreted by the authors as an example of the self-defining "fallacy of the last move". A 1987 report from a study group of the American Physical Society concluded after more than a year with full access to

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Artistic licence — a space-based railgun for interception of ballistic missiles.

SDI technology that "... a decade or more of intensive research would be required to provide the technical knowledge needed for an informed decision about the effectiveness and survivability of directed-energy weapon systems".

Of the two SDI options now under construction, one is the old two-layer terminal defence, in which the exo-atmospheric intercepts are limited by the continued inability to discriminate warheads from decoys in space (when the reentry vehicles carrying the warheads will have used 'antisimulation' to enable them to resemble cheap decoys), and the endo-atmospheric intercepts continue to suffer from high cost of interceptor, local exhaustion and penetration techniques.

The other option, a largely space-based defence consisting of space-based interceptors (SBI), has problems of survivability as well as of performance. Lakoff and York discuss the current SBI strategy, in which thousands of "brilliant pebbles", each weighing about 50 kg, would lurk in

space as individual interceptors in their "life jackets". According to a Pentagon news briefing given on 9 February 1990 by Lt Gen. George Monahan and the originator of the brilliant-pebbles concept, Dr Lowell Wood of the Lawrence Livermore National Laboratory, 4,614 pebbles would be built at a cost of \$1.4 million each. This contrasts with Wood's previous assertion that brilliant pebbles could be built and launched for a total cost of \$20,000 each and that the system "could be fully shared with the Soviet Union with minimal complications for national security", in view of the claimed "off-the-shelf

technology" incorporated into the system. But such a system prescribes its own counter-system; in every respect the anti-satellite weapon that could be used to shoot down each brilliant pebble within days after its launch to orbit is simpler, cheaper and less challenging than the SBI itself.

First, consider launch cost. An ASAT homing head could be lofted to orbital altitude with a rocket system about five times lighter than would be required to launch its payload into orbit. Second, consider life. The largely autonomous SBI needs a life in orbit of five to ten years; the ASAT needs a life of ten minutes. ASAT

would be commanded by ground-based radar, which can be used over and over again as individual ASATs destroy individual SBIs. Finally, consider sensor performance. The SBI may need to look out 1,000 km with a broad field of view; the ASAT need look no farther than 20 km to find its assigned prey, with a field of view that is little more than a peephole. If the SBI system manages to intercept the ASAT before the ASAT can complete its intercept, an SBI destroying an ASAT is as dead as one that has been destroyed by the ASAT.

How do proponents maintain their enthusiasm for space-based interceptors that can be shot down as they are being launched much more cheaply than it cost to put them up? By invoking the evolution of the Russian bear into the Soviet pussycat, which can now be relied upon to allow weapons based in space in violation of the 1972 US-Soviet anti-ballistic missile treaty (or after its denunciation) freely to pass over Soviet territory! I believe that

Science Photo Library

neither the Soviet Union nor the United States would permit each other's vulnerable weapons to survive; nor should they.

The subject of *A Shield in Space?* is complicated, as is the organization of the book. The book pleases by the inclusion of a heading at the top of each of the 30 pages of notes, citing the corresponding page range of the text. Its 11 pages of bibliography, 11 pages of index, and 5 pages of acronyms and abbreviations are also well done. But the authors miss the mark in not including a single cartoon of a candidate layered defence system, and in citing the initial proposal for \$26 thousand million for the first five years of a ten-year research programme without the accompanying tab of \$70 thousand million for the full ten-year programme.

In a fascinating portion of the book, the authors quote Robert C. McFarlane, the national security advisor to Reagan at the time of the 1983 "Star Wars" speech and the person responsible for developing and concealing the SDI programme until its announcement. In his testimony to Congress in 1988, McFarlane recounted how Reagan personally changed McFarlane's and the Joint Chief of Staff's SDI concept from one of protecting missiles to an essentially perfect defence of the US public. McFarlane, as a government official, had argued for the SDI programme in background interviews and discussions with individuals inside government and out as "the greatest sting operation in history", rather than as an effective system. And in his testimony, he noted how profoundly antidemocratic government has become in pursuit of support for its policies. "I am not sure how it would come out, but I do know that we have never had the integrity to explain it in those terms." *A Shield in Space?* weighs in on the side of that missing integrity. □

Richard L. Garwin is at the IBM Thomas J. Watson Research Center, PO Box 218, Yorktown Heights, New York 10598, USA.

■ In *New Strategy Through Space*, just published by Leicester University Press, Neville Brown discusses the future of space research. He argues that, since the announcement of SDI in 1983, the debate has focused too narrowly on military issues, and that space-based defence is a destabilizing influence on world politics. He puts forward a scheme of planetary management using space-based surveillance and telecommunications. Price is £35.

■ *Ways Out of the Arms Race*, edited by J. Hassard *et al.*, subtitled *From the Nuclear Threat to Mutual Security*, reports the proceedings of the Second International Scientists' Congress (1988). Published by World Scientific, price £23, \$36.

■ W. A. Schwartz and C. Derber's *The Nuclear Seduction* considers conventional wisdom about nuclear war to be profoundly misleading while the main forces endangering the world remain largely invisible. Published by the University of California Press, price \$25.

Ignorant genes

J. Bruce Walsh

The Wisdom of the Genes: New Pathways In Evolution. By Christopher Wills. *Basic Books*: 1989. Pp. 351. \$19.95.

THE flood of molecular data has drenched evolutionary biology as much as any field, and has initiated often heated debates as to the function (if any) of particular aspects of genome organization. Although there is no doubt that features such as mobile genetic elements and the intron/exon structures of many genes have important evolutionary consequences, the nature of the evolutionary forces responsible for the origin and subsequent maintenance of these structures is contentious. One school argues that these features are maintained because they facilitate the evolutionary process. Proponents of this view generally, implicitly or explicitly, invoke species selection (the differential formation and extinction of species), a notion that makes many evolutionary biologists uneasy. The alternative school holds that the potential for evolutionary facilitation has little bearing on origin and maintenance. For example, although some claim that the ability of mobile elements to generate mutations is sufficient to account for their ubiquity, these elements can spread through a population even if deleterious. So even though the insertion or deletion of a particular element can result in a favourable new mutant, mobile elements would still be expected to be widespread even if no such mutants were ever generated.

Wills has written an engaging popular book examining these issues, arguing that previous evolution has structured the genome in such a way as to expedite future evolution. As an introduction into modern evolutionary biology for the general reader, his book is a success, capturing the excitement and breadth of the subject. The main strength of the book is its wealth of amusing personal and historical anecdotes that make for enjoyable reading. Its scope is also fascinating, ranging over such seemingly random topics as the similarity between the way a tuxedo shop operates and the organization of antibody genes, müllerian mimicry rings in tropical butterflies, the Stanley Steamer versus Henry Ford assembly lines, and mobile genetic elements.

Wills is less successful in arguing that evolution indeed does "work better" now than in the past. Acknowledging the problems in invoking species selection, Wills suggests that structures facilitating evolution can be maintained by selection on individuals. His strongest case is for mimicry in butterflies, in which there are probably selection pressures constantly to

change wing patterns over short intervals of evolutionary time. A genetic organization facilitating rapid changing in wing patterns, Wills argues, will be favoured by individual selection. That certain characters have been selected to be developmentally labile is reasonable. What is unclear is how widespread such characters are and how this developmental flexibility is achieved. Wills seems to be arguing that fairly large-scale changes are required, along the lines of constructing a new gene as opposed to making a few changes within an existing gene. In my opinion, Wills does not make a compelling case for this idea. It seems that genes are neither wise nor stupid with respect to their future evolution, just ignorant. □

J. Bruce Walsh is in the Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona 85721, USA.

Dangerous stuff

Alastair Hay

Halogenated Biphenyls, Terphenyls, Naphthalenes, Dibenzodioxins and Related Products, 2nd edn. Edited by R. D. Kimbrough and A. A. Jensen. *Elsevier*: 1989. Pp. 518. Dfl. 375, \$192.50, £120.

SHIPS bringing hazardous waste to Britain for disposal were welcome a year ago. They are not greeted so rapturously now, particularly when they feature as the lead story in national news bulletins. Last summer, Britain's traffic in hazardous waste almost ground to a halt. Ships with a cargo of waste for disposal in the United Kingdom were prevented from docking both by the harrying activities of environmentalists and by the dockers themselves, who refused to unload the cargo.

The principal target of the protesters was consignments of polychlorinated biphenyls (PCBs) which were to be destroyed by burning in an incinerator in Wales. Despite reassurances from the government that incineration of the waste presented no risk, local residents argued to the contrary, winning the support of the dockers.

With no chance of offloading their cargo, the shippers had to return it to the sender, in this case the Canadian state government of Ontario. The British Government also had egg on its face over the episode. What it had seen as a highly profitable trade — an environmental service, so to speak — was viewed by many others as yet more evidence that Britain was becoming the dustbin of Europe. Fewer cargoes of waste are likely to come to Britain in future.

This campaign against PCBs, spearheaded by the environmental group

Greenpeace, brought the issue of hazardous waste disposal in Britain into the limelight. Questions raised at the time are still largely unanswered, yet are still relevant. Should PCBs be destroyed by incineration? If so, where? Can this be done safely? Regrettably, there are no simple answers. But anyone wrestling with these issues will find it helpful to consult Kimbrough and Jensen's superbly edited book. The title is enough to deter any but the most intrepid of readers, but

transformers, how can they be destroyed? Evidence suggests that this can be done reasonably well in some, but not all, incinerators. It is this variability between furnaces that makes people who live near them apprehensive. The alternative of not destroying the chemicals is even more unattractive; this book shows just how much environmental damage PCBs have caused.

One concern about PCBs and dioxins is their carcinogenic potential. Many investigators have tried to assess this in occupa-

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Chilly reception — the *Ndezhea Obukhova*, with an unpopular cargo from Canada.

the publishers, clearly pleased with the reception given to an earlier edition published 9 years ago, decided to play safe and keep a familiar label.

Chapters written by different authors deal with the production, chemical, physical and toxicological properties of PCBs and their brominated cousins. Chlorinated and brominated dibenzodioxins (dioxins) and dibenzofurans (furans) produced during the incineration of halogenated biphenyls are similarly covered. The book is appealing because each author looks at some aspect of all these chemicals, noting similarities where they occur and pointing out structural relationships that explain, for example, the cytochrome P_{450} enzyme-inducing properties of particular PCBs or dioxins.

For regulators and those concerned about the environmental impact of PCBs, dioxins and furans, the book contains considerable food for thought. Because of their insulating properties, PCBs were used in transformers and capacitors by the electricity-supply industry. With the death of 15,000 guillemots in the Irish Sea in 1969 and the declining fertility of many populations of sea mammals in the 1970s blamed on the PCBs, manufacturers in the West decided to stop production (one chapter suggests that PCBs may still be being produced in developing countries). Alternatives are now being used, but it will take until the year 2000 before most of the tens of thousands of tons of PCBs still in use are replaced.

Once PCBs have been drained out of

tional studies. The delay between writing and publication of the book means some occupational information about exposure to chlorinated dibenzodioxins and the herbicides 2,4,5-trichlorophenoxyacetic acid and 2,4 dichlorophenoxyacetic acid is not included. The carcinogenic risk to man of exposure to these chemicals is still hotly debated, with evidence from Sweden (over a 10-year period) suggesting that there is an increased risk of soft-tissue sarcoma and malignant lymphoma in those exposed through their occupation. Studies conducted elsewhere both support and question the Swedish findings, and it can only be said at this stage that the jury is still out. One study of 121 workers exposed to the most toxic of the dioxins at a Monsanto chemical plant in 1949 is cited here as showing that there were less cancers than expected in the exposed workers. This study has since been discredited and should thus be disregarded. Regrettably, the authors do not refer to the flaws in the study even though this information has been in the public domain for several years.

But these are minor faults in an extremely valuable book. Purchasers may wince at the price and be disappointed that the book contains no easy solution. Ultimately, however, I suspect that readers will treat it rather like a good bath and luxuriate in what it has to offer. □

Alastair Hay is in the Department of Chemical Pathology, University of Leeds, Leeds LS2 9JT, UK.

Getting into the swing

B. J. Hiley

The Physics of Vibration: Omnibus Edition. By A. B. Pippard. Cambridge University Press: 1989. Pp.639. £35, \$65.

THIS omnibus edition amalgamates two previously published volumes: part 1 contains an exposition of the classical theory of vibrations; part 2 extends the discussion to their quantum-mechanical aspects. Much of the content remains the same as that published previously, but the author has made minor additions and rectified the errors that inevitably occur in a compendium of this size.

As I did not study the original volumes in detail when they first appeared, I am now reporting my first impressions. Overall, I found the presentation very refreshing. Pippard takes great care to explain all the physical principles underlying the phenomena he examines, not allowing mathematics to dominate the discussions and explaining fully the physical implications of the formulae. The obvious advantage of this approach is that as the system becomes more complicated and an exact mathematical treatment more difficult — that is, the system is part of the real world — the necessary insights are at hand to penetrate the essential features of the system without the need to become too entangled in the mathematics. It is this approach that makes the book attractive, not only to the specialist, but also to physicists and engineers in general.

Pippard's overall strategy is to start with a discussion of the simplest of systems and then to examine the change of behaviour as more and more complications are added. For example, the simple harmonic oscillator is first modified by adding anharmonic terms, then dissipation is included. The topic is further generalized by introducing various types of driving forces, and so on. Every development is illustrated with copious examples, which include mechanical systems, a.c. circuits, various types of electronic circuits, and atomic and molecular systems. This free use of examples from different branches of physics illustrates the universality of oscillator-type problems.

The extension to quantum-mechanical systems presents a much more difficult problem and requires a more detailed mathematical approach. It is not Pippard's intention to give a full quantum-mechanical treatment of these systems, rather his aim is to show that in many situations the quantum-mechanical solutions turn out to be very close to those obtained from classical analyses. This has two advan-

tages. First, in many quantum-mechanical situations, the classical results can be used as a guide to thinking about the system under consideration. Second, quantum-mechanical calculations from first principles are often daunting, whereas the classical problem is more tractable. Of course, there is a disadvantage to this whole approach, of which Pippard is very aware — quantum-mechanical systems do produce effects that cannot be understood in terms of classical physics. Therefore great care must be taken when using the classical results. Indeed, Pippard's aim is to generate sufficient feeling for the subject so that such pitfalls can be avoided.

My only reservations arose in the last part of the book where I felt the reader was asked to accept too many statements without adequate explanation or justification. This directly contrasts with the first part, where no such criticism can be made, and could possibly have arisen from the nature of the task that was being attempted.

The text contains many fine illustrations and diagrams which considerably enhance the explanations. To sum up, then, this is an excellent treatment of the subject. □

B. J. Hiley is in the Department of Physics, Birkbeck College, Malet Street, London WC1E 7HX, UK.

Global view

Bob Chapman

The World at 18 000 BP, Volume 1: High Latitudes. Volume 2: Low Latitudes. Edited by Olga Soffer and Clive Gamble. Unwin Hyman: 1989. Pp.353; 344. Each volume £45, \$65.

WHAT was life like 18,000 years ago, and why should we be interested? These questions are central to current research on human and cultural evolution, and on variability among hunters and gatherers past and present. A perennial issue is that of elucidating the relationships between environmental and cultural change. In the case of hunters and gatherers, can we progress from the often static pictures drawn from ethnography and tease apart the variables that interacted to cause the evolution and diversity of cultures throughout the world? At what scales should these variables be measured, and how can they be analysed using archaeological data? These questions go right to the heart of archaeology as a discipline, as they address our ability to understand the archaeological record.

The world 18,000 years before present (BP) was experiencing the last glacial maximum (LGM), during which time most of the continental ice sheets reached their maximum extension, and temperatures were at their minimum values. How did Palaeolithic hunters and gatherers cope with these extreme conditions of climate? What was the adaptive potential of developments in technology, or settlement and subsistence strategies, or social organization? How far did changes in demography and culture depend on geographical considerations? Can we make any inferences about the effects of these changes among hunters and gatherers upon their gene flow and their local evolutionary trajectories?

Olga Soffer and Clive Gamble conceived the LGM project to study these problems. This, the project's first publication, provides syntheses of environmental

and cultural change in seven areas of the world. Coverage of local regions varies according to the current state of research (controversy over the antiquity of human colonization of the New World led the editors to limit discussion to one chapter) and, presumably, to the availability of authors. The result is quite the best modern synthesis of late glacial hunters and gatherers in their global context. Apart from any other consideration, the book finally routs the Eurocentric bias present in Upper Palaeolithic research for more than a century.

What can be learnt from the LGM project, and what problems can be identified for the future? Soffer and Gamble argue that during the LGM some areas were depopulated whereas in others humans retreated into 'refuges', such as in south-west France/northern Spain and the Russian plain. After the LGM, many areas became repopulated (for example, northern Europe and the arid parts of Africa). That this is not some simple outcome of environmental determinism is shown by the observation that people lived within 250 kilometres of the ice sheet of the Russian plain during the LGM. The dangers of determinism are further emphasized by Wobst, who stresses that in the LGM, humans were "the most widely distributed mammal worldwide", that they were "at the top or close to the top of every regional food chain" and that they were "important in restructuring food chains". The seemingly independent changes of technology in relation to environment which occurred in low-latitude areas such as southern Africa thus merit further thought.

The differences in environment and cultural change between high/northern and low/southern latitudes during the LGM are made apparent by global study. The editors and Wobst point out "the relative uniqueness of the periglacial behavioural package (such as portable 'art', cave painting, rapid change in lithic assemblage types, narrow stylistic standards, lots of concern for the non-utilitarian features of tools), compared

with the rest of the world". Such a view makes well-studied areas like the Dordogne no less important, but puts them within a wider context of variability among hunters and gatherers. Such variability can usefully be compared with what is visible in the ethnographic record.

Of course, much depends on our ability to reconstruct environments during the LGM, and to relate their changes to those of hunter-gatherer cultures in different parts of the world. Here, van Andel offers wise council about the scale and temporal resolution of palaeoenvironmental and archaeological data. He points to the "highly generalised" nature of palaeoenvironmental maps, to the inability to "deal with rare or catastrophic events" and to the "impossibility of firmly establishing regional or continent-wide synchronicities between palaeoenvironmental and archaeological events".

But how reliable is our knowledge of the archaeological record in different parts of the globe? This question directly affects one of the central aims of the LGM project, that of comparability. Biases in surveys clearly exist (they are more difficult in equatorial rainforest than in denuded regions, for example), and the taxonomic organization of mainly lithic materials can determine one's perception of change in human occupation. Geomorphological and other changes in the Earth's surface can remove large areas from archaeological visibility. In areas such as Epirus in Greece, it is argued that the visible archaeological record may have been peripheral to the area of main settlement and population distribution in the LGM. Taking this one step further, our knowledge of contemporary hunters and gatherers leads us to realize that changes in their subsistence and mobility strategies (see the chapter by Weniger on Germany), as well as in their technologies (such as the transition from flake to blade production, discussed by Edwards), can determine the visibility and intensity of the archaeological record. Thus, inferences on the continuous or discontinuous presence of human populations in different areas of the world during the LGM require careful evaluation. It is also clear from these volumes that hidden assumptions about the meaning of variation in lithic assemblages, whether demographic or ethnic, form one area for discussion.

Soffer and Gamble are correct in their assertion that the LGM project "is an exercise in a new approach to world prehistory". The parameters have been defined, contributors organized, the advantages demonstrated and the problems identified. Upper Palaeolithic studies will never be the same again. □

Bob Chapman is in the Department of Archaeology, University of Reading, Reading RG6 2AA, UK.

Collision-induced two-ion chaos

R. G. Brewer*, J. Hoffnagle*, R. G. DeVoe*, L. Reyna† & W. Henshaw†

* IBM Research Division, Almaden Research Center, 650 Harry Road, San Jose, California 95120-6099, USA

† IBM Research Division, T. J. Watson Research Center, Yorktown Heights, New York 10598, USA

Two ions confined to a radio-frequency trap and cooled by radiation pressure exhibit deterministic chaos. Theoretical analysis of the two-ion dynamics reveals that the route to chaos is due solely to ion-ion collisions. During a collision, the nonlinear Coulomb interaction introduces a transient instability, which gives way to stable single-particle-like motion as the ions move apart. The chaotic dynamics are characterized by a strange attractor that resembles a spiral galaxy.

POINCARÉ found that, for the three-body problem of celestial mechanics¹, the solutions to the equations of motion (which are nonlinear because of gravity) display an unusual richness and complexity. We now know that the system represents an example of deterministic chaos. A requirement for chaos is nonlinearity in the equations of motion, which makes them non-integrable and necessitates numerical solution. Experimental tests of such computer calculations have been restricted largely to macroscopic systems, such as turbulence in liquids. Although few-body systems are of fundamental interest, the only example studied so far of two-body chaos on the atomic scale seems to be the hydrogen atom². Here we consider another example, the case of two trapped ions³.

Wuerker *et al.*⁴ observed a phase transition when charged aluminium particles were stored in a Paul or radio-frequency (r.f.) electric quadrupole trap. The aluminium particles assumed an ordered (crystalline) state when cooled by collisions with a background gas, but adopted a disordered state, characterized by a random motion with large orbits, when a control parameter (the r.f. trap voltage) achieved a critical value. The same phenomenon was observed recently with small numbers of trapped ions cooled by the radiation pressure of a laser beam^{3,5}. We speculated that the ion-trap system must be deterministic rather than being dominated by stochastic forces, and therefore that order-to-chaos transitions were being observed in which the nonlinearity arises from the ion-ion Coulomb interaction. Indeed, an experimental and theoretical study of the simplest case of two trapped Ba⁺ ions revealed a transition to chaos³.

Nevertheless, the route to two-ion chaos has been puzzling, because standard scenarios^{6,7} such as a cascade of period doublings or intermittency do not seem to apply. Here we propose that ion-ion collisions sustain the two-ion chaotic state, resulting in a new type of strange attractor. As the ion-ion Coulomb interaction introduces the nonlinearity that leads to chaos, the importance of collisions is perhaps unsurprising. We find that in the chaotic state, the ions are sufficiently energetic to undergo close encounters or hard collisions, which trigger 'bursts' of nonlinearity that dominate the dynamics. During these collisions the ions absorb r.f. energy, being balanced at equilibrium by laser damping. But as the ions move far apart, the nonlinearity is essentially turned off and the relative motion takes on a single-particle-like character.

One of the key signatures of chaos is an extreme sensitivity to initial conditions. Two neighbouring orbits in a chaotic region of phase space will diverge exponentially with a rate of divergence characterized by the Lyapunov exponent^{6,7}. For two non-rotating ions, the relative motion takes place in a four-

dimensional phase space (two position coordinates and two momenta), so four Lyapunov exponents exist, one for each dimension. As the ions are damped by laser cooling, at least one of the exponents must be negative, and if the ions are chaotic, at least one exponent must be positive. Our analytic and numerical calculations of the Lyapunov exponents indicate that the positive exponent is due solely to collisions. The important issue of how the initial conditions necessary for chaos are achieved will be reported elsewhere.

Analytic stability analysis

We consider the motion of a pair of ions in a Paul trap with cylindrical symmetry and an oscillatory quadrupole potential $V = V_{ac}(\cos \Omega t)(z^2 - (x^2 + y^2)/2)/\bar{r}_0^2$ where $\bar{r}_0$ is the trap radius and V_{ac} is the amplitude of the applied r.f. potential with frequency Ω . Introducing dimensionless parameters for time, $\tau = \Omega t/2$, and potential, $q_3 = 4eV_{ac}/(m\Omega^2\bar{r}_0^2)$, where e and m are the ion's charge and mass, one finds that the equations of motion separate into centre-of-mass and relative coordinates. The former obey a Mathieu equation⁸⁻¹¹ with well-known analytic solutions, whereas the relative coordinate R obeys a modified Mathieu equation³

$$\ddot{R}_i + \Gamma \dot{R}_i + [2q_i \cos 2\tau - \alpha/R^3]R_i = 0 \quad (1)$$

where Γ is a laser damping parameter, $\alpha = e^2/(m\Omega^2)$ is a Coulomb coupling parameter and the dots denote differentiation with respect to τ . Here, $q_1 = q_2 = -q_3/2$ and the subscripts $i = 1, 2, 3$ designate the i th cartesian component of R . Equation (1) is a set of three coupled equations with three degrees of freedom in a six-dimensional phase space. One degree of freedom in equation (1) is simply a conserved angular momentum which does not contribute to the chaotic dynamics. We have performed calculations both with the full three-dimensional version of equation (1) and with a reduced set of two equations having only one radial and one axial coordinate (two degrees of freedom in a four-dimensional phase space) corresponding to the case of zero angular momentum; we find essentially the same results in both cases. In the ordered state, the ions are locked into

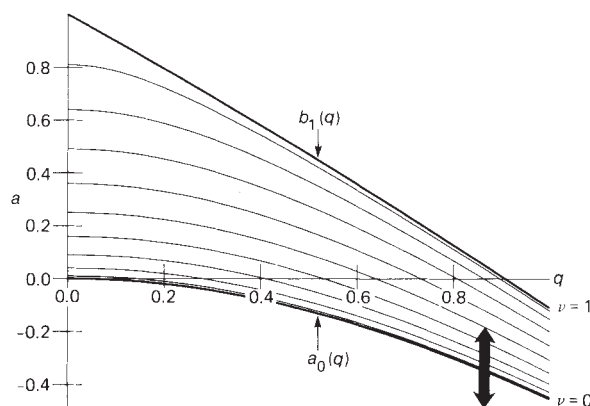


FIG. 1 Plot of a against q for the Mathieu equation¹¹, where the two-ion Coulomb interaction simulates the a parameter. The first stability region for the motion along the z coordinate falls between the boundaries $a_0(q)$ and $b_1(q)$; the motion is unstable elsewhere. The heavy arrow marks the region of transient two-ion instability at $q = 0.86$. The Floquet exponents $0 \leq \nu \leq 1$ are shown, and for clarity, the stability region for the r motion is not.

TABLE 1 Numerically and analytically derived Lyapunov exponents

	λ_1	λ_2	λ_3	λ_4	$\sum_{i=1}^4 \lambda_i$
Numerical	0.008675	0.001560	-0.001758	-0.008885	-0.000408
Analytic (equation (14))	0.007915	-0.000100	-0.000100	-0.008113	-0.000398

The damping constant $\Gamma = 0.0002$; equation (11) predicts $\sum_{i=1}^4 \lambda_i = -2\Gamma$.

position, executing only small harmonic oscillations—the fast r.f. or ‘micro’-motion and the slow secular motion caused by the effective (time-averaged) r.f. potential. Increasing the r.f. field to a critical value ($q \approx q_3 = 0.86$) induces the order-to-chaos transition, whereupon the two ions move in large erratic orbits while still being bound in the trap.

Instability. When $q \geq 0.91$, solutions of equation (1) show that the ions absorb r.f. energy until they strike the trap walls and thus are lost; in other words, the trap becomes unstable. This is the well-known prediction of the first stability region of the Mathieu equation (equation (1) with $\Gamma = 0$ and $-\alpha/R^3 \rightarrow a$) and is shown in Fig. 1 as a plot of a (defined in ref. 11) against q , where the lines $b_1(q)$ and $a_0(q)$ bound the stable region¹¹. For the chaotic region $0.86 \leq q < 0.91$, we show below that an ion-ion collision generates a transient instability (heavy arrow of Fig. 1), with the ions absorbing r.f. energy over the short collision period. On average, the increase in energy is just balanced by the damping that results from laser cooling, and thus the chaotic ions remain bound.

Modified Mathieu equation. For analytic calculations, we assume a four-dimensional set of equations

$$\dot{\mathbf{x}} = f(\mathbf{x}) \quad (2)$$

where the variation in two neighbouring trajectories obeys

$$\delta \dot{\mathbf{x}} = \frac{\partial f}{\partial \mathbf{x}} \delta \mathbf{x} \quad (3)$$

and $\partial f/\partial \mathbf{x}$ is the jacobian matrix whose eigenvalues characterize the orbits^{6,7}. We can show that λ_1 , the largest of the four Lyapunov exponents and the one responsible for chaos, can be estimated by the time-average of the largest eigenvalue of the jacobian,

$$\lambda_i = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \text{Re} \int_0^\tau \left(\frac{\partial f}{\partial \mathbf{x}} \right)_{ii} d\tau' \equiv \left\langle \left(\frac{\partial f}{\partial \mathbf{x}} \right)_{ii} \right\rangle \quad (4)$$

the other three exponents being less certain because the vector $\delta \mathbf{x}$ changes direction with time. λ_i is determined by the real part of $\partial f/\partial \mathbf{x}$, because the imaginary part, corresponding to an oscillation, makes no contribution. In the chaotic state, the amplitude of any small variation $\delta \mathbf{x}(0)$ grows, on average, exponentially in time as $e^{\lambda_1 \tau}$.

The non-autonomous equations of motion (1) can be simplified by recognizing that the time of a hard collision is the fastest timescale in the problem, followed by the micro-motion period, the secular period and the damping time, in that order; the ratios are approximately 0.001:0.3:2:500. During a collision, the four-dimensional form of equation (1) reduces to Kepler's equations of motion¹⁰:

$$\ddot{r} - r\alpha/R^3 = 0 \quad (5a)$$

$$\ddot{z} - z\alpha/R^3 = 0 \quad (5b)$$

with $r = \sqrt{x^2 + y^2}$ and $R = \sqrt{z^2 + r^2}$ and where the damping and micro-motion terms are negligible. Pre-empting the final result in equation (14), we find that equation (5) alone gives the essential contribution to the Lyapunov exponents through the time-averaged Coulomb term, which peaks sharply during a collision, and it is this quantity that can give a positive Lyapunov exponent and chaos. The time between collisions yields the damping term Γ in equation (14). Between collisions, a perturba-

tion theory that extends that of Kapitza⁹ yields the secular equations of motion,

$$\ddot{r} + \Gamma \dot{r} + r \left(-\frac{\alpha}{R^3} + \frac{q^2}{8} \right) = 0 \quad (6a)$$

$$\ddot{z} + \Gamma \dot{z} + z \left(-\frac{\alpha}{R^3} + \frac{q^2}{2} \right) = 0 \quad (6b)$$

because the micro-motion is faster and can be separated in the usual way^{9,10}. Equation (6) retains the Coulomb terms, and so can apply to the collision periods as well as to the collision-free intervals, where now only the micro-motion has been ignored. Defining $\rho = \dot{z}$ and $\beta = \dot{r}$, we introduce a four-dimensional vector

$$\mathbf{x} = \begin{pmatrix} z \\ \rho \\ r \\ \beta \end{pmatrix}$$

so that equation (6) reduces to four first-order equations, as in equation (2):

$$\begin{aligned} \dot{z} &= \rho \\ \dot{\rho} &= -\Gamma \rho - z \left(-\frac{\alpha}{R^3} + \frac{q^2}{2} \right) \\ \dot{r} &= \beta \\ \dot{\beta} &= -\Gamma \beta - r \left(-\frac{\alpha}{R^3} + \frac{q^2}{8} \right) \end{aligned} \quad (7)$$

The resulting determinant of the jacobian takes the form

$$\begin{vmatrix} -\lambda & 1 & 0 & 0 \\ J_{10} & -(\Gamma + \lambda) & J_{12} & 0 \\ 0 & 0 & -\lambda & 1 \\ J_{30} & 0 & J_{32} & -(\Gamma + \lambda) \end{vmatrix} = 0 \quad (8)$$

with $J_{10} = \alpha/R^3 - q^2/2 - 3\alpha z^2/R^5$, $J_{12} = J_{30} = -3\alpha rz/R^5$, and $J_{32} = \alpha/R^3 - q^2/8 - 3\alpha r^2/R^5$. The Lyapunov exponents, which are the time-average of the eigenvalues of equation (8) become

$$\lambda_i = \text{Re} \frac{1}{2} \langle -\Gamma \pm \sqrt{\Gamma^2 + 4\theta_{1,2}(\tau)} \rangle; \quad i = 1 \text{ to } 4 \quad (9)$$

where

$$\begin{aligned} \theta_{1,2}(\tau) &= \frac{1}{2} \left\{ -\left(\frac{\alpha}{R^3(\tau)} + \frac{5}{8} q^2 \right) \right. \\ &\quad \left. \pm 3 \left[\left(\frac{\alpha}{R^3(\tau)} \right)^2 + \alpha \frac{(z^2(\tau) - r^2(\tau))}{4R^5(\tau)} q^2 + \frac{q^4}{64} \right]^{1/2} \right\} \end{aligned} \quad (10)$$

An immediate consequence of equation (9) is that it obeys the sum rule

$$\sum_{i=1}^4 \lambda_i = -n\Gamma \quad (11)$$

where $n = 2$ is the number of degrees of freedom in the two-ion problem. It is easy to show that equation (11) expresses the rate of change of the volume in phase space, $(1/V)(dV/d\tau) = -2\Gamma$, and thus vanishes when $\Gamma = 0$ (Liouville's theorem), as is true for any hamiltonian system, including the non-dissipative two-ion trap system. To evaluate the individual Lyapunov exponents predicted by equation (9), we followed the trajectory $r(t)$, $z(t)$ of equation (1) in the chaotic state for a given initial condition. The resulting time dependence of the argument of equation (9), $\Lambda = -\Gamma \pm \sqrt{\Gamma^2 + 4\theta_{1,2}(\tau)}$, is shown in Fig. 2, and the time-average of this behaviour yields the four Lyapunov exponents listed in Table 1, these being independent of the initial conditions as long as chaos is induced. The role of hard collisions is now apparent: they are responsible for the sharp features of Fig. 2a, whereas between collisions the contributions on this scale are not discernible.

It is clear from equation (9) that the ordered state will be stable whenever $\sqrt{\Gamma^2 + 4\theta_{1,2}}$ is imaginary, and unstable when it is real. To proceed further, we recognize that the ion-ion distance in the ordered state is dominated by the equilibrium radial separation and not by small oscillations, so that $r^2 \gg z^2$ and equation (10) becomes

$$\theta_1(\tau) = \frac{\alpha}{R^3(\tau)} - \frac{q^2}{2} \quad (12a)$$

$$\theta_2(\tau) = -\left(\frac{2\alpha}{R^3(\tau)} + \frac{q^2}{8}\right) \quad (12b)$$

In equation (12b), $\theta_2 < 0$ holds for any R , but equation (12a) implies that a critical value of R divides the stable and unstable regions, that is, instability occurs when

$$\frac{\alpha}{R^3(\tau)} > \frac{q^2}{2} \quad (13)$$

Equation (13) is to be compared to the Mathieu equation instability condition¹¹ $-a > q^2/2$; here the time-dependent Coulomb interaction $-\alpha/R(\tau)^3$ corresponds to the static destabilizing d.c. field (the a term). A collision satisfying equation (13) can be viewed as a dynamic or transient instability in the

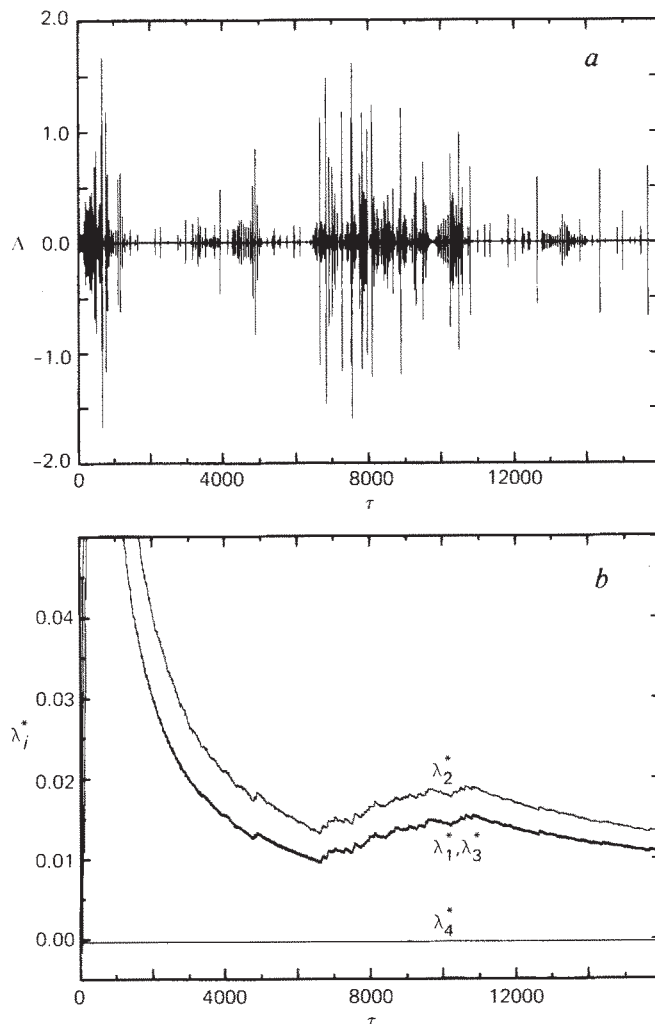


FIG. 2 *a*, Time-dependent behaviour of $\Lambda = -\Gamma \pm \sqrt{\Gamma^2 + 4\theta_{1,2}(\tau)}$, the argument of equation (9), evaluated in time bins of $\Delta\tau=2$ for an r, z trajectory derived numerically from equation (1). The sharp features are the result of hard collisions. The initial conditions are $(r_0, z_0, \dot{r}_0, \dot{z}_0) = (5.88, 0.3, 0.1, 0.1)$, $q=0.86$ and $\Gamma=0.0002$. *b*, Numerical solutions of λ_{1-4}^* as a function of time. There is an increase in the interval 7,000–11,000, where the density of collisions (see *a*) is large, followed by convergence. Conditions are the same as in *a*.

following sense. During a collision, the Coulomb interaction generates a transition, at a given q , across the lower instability boundary $a_0(q)$ of Fig. 1. As the ions rebound from the collision and move far apart, the nonlinearity is turned off, causing the ions to return to their stable configuration in which their relative motion has essentially single-particle character. The transitions back and forth follow a random sequence of collisions dictated by the chaotic nature of the orbits. In this way, the two-ion trap system functions as a relaxation oscillator. Note that the upper instability boundary $b_1(q)$ of Fig. 1 is irrelevant for these arguments (compare the discussion of ref. 12).

One might conclude from these considerations that chaos can occur at any q value. On the contrary, recent experiments³ and further work of our own (manuscript in preparation) show that the onset of two-ion chaos appears at a specific value, namely $q \approx 0.86$. Although equation (13) can be satisfied at any q value by a sufficiently energetic initial condition, the time-average behaviour of equation (9) will determine whether the chaotic behaviour is stationary or transient. Thus equation (13) seems to be a necessary but not a sufficient condition for stationary chaos.

Between collisions, the ions are sufficiently far apart most of the time that the q^2 terms in equation (12) dominate, with $\theta_1 = -q^2/2$ and $\theta_2 = -q^2/8$, the contribution to the Lyapunov exponents being $\lambda_{1-4} \rightarrow -\Gamma/2$. But during a collision the q^2 terms can be neglected, so that $\theta_1(\tau) = \alpha/R(\tau)^3$, $\theta_2(\tau) = -2\alpha/R(\tau)^3$, and the contributions in these intervals are $\lambda_{1,4} \rightarrow \pm\sqrt{\alpha/R(\tau)^3}$ and $\lambda_{2,3} \rightarrow -\Gamma/2$. The time-average behaviour then assumes the form

$$\begin{aligned} \lambda_1 &\approx -\Gamma/2 + \langle \sqrt{\alpha/R(\tau)^3} \rangle, \\ \lambda_2 &= \lambda_3 = -\Gamma/2, \\ \lambda_4 &\approx -\Gamma/2 - \langle \sqrt{\alpha/R(\tau)^3} \rangle. \end{aligned} \quad (14)$$

For a typical hard collision, $\sqrt{\alpha/R(\tau)^3} \approx 100$. Assuming from Table 1 that $\lambda_1 = 0.008$, $\Gamma = 0.0002$ and $\langle \sqrt{\alpha/R(\tau)^3} \rangle \approx P \times 100$, we see that the probability for a collision is of the order of $P = 0.0001$, so that 99.99% of the time the ions are essentially single-particle-like. Nevertheless, the small time that the collisions are active is sufficient to give rise to the positive Lyapunov exponent responsible for chaos.

Numerical solutions

As a check on the semi-analytic results above, we have calculated the Lyapunov exponents numerically for the chaotic state using the technique of Shimada and Nagashima¹³. For the four-dimensional flow expressed by equation (1), we follow the time

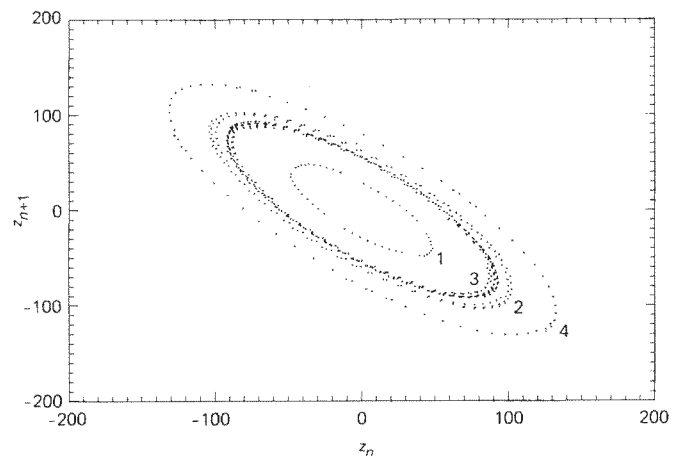


FIG. 3 Return maps of the chaotic state. Here z_{n+1} is plotted against z_n strobed at the micro-motion period. Transitions from one orbit to another are the result of collisions. Time advances as the orbit number increases. Here $(r_0, z_0, \dot{r}_0, \dot{z}_0) = (75.674, -11.034, 40.391, -6.947)$, $q=0.86$ and $\Gamma=0.0002$.

evolution of a reference trajectory, $\mathbf{x}$, and four perturbed trajectories that differ initially by $\delta\mathbf{x}_i$. The vectors $\delta\mathbf{x}_i$ are orthogonal and of length δx_0 . After some small time interval $\Delta\tau$ the perturbed and reference trajectories differ by $\mathbf{h}_i(\Delta\tau)$:

$$\begin{aligned}\mathbf{x}(0) &\rightarrow \mathbf{x}(\Delta\tau) \\ \mathbf{x}(0) + \delta\mathbf{x}_i &\rightarrow \mathbf{x}_i(\Delta\tau) + \mathbf{h}_i(\Delta\tau).\end{aligned}\quad (15)$$

Consider the growth of the N -dimensional phase-space volume defined by the vectors $\mathbf{h}_i(\Delta\tau)$ to $\mathbf{h}_N(\Delta\tau)$:

$$V_N(\Delta\tau) = |\mathbf{h}_1(\Delta\tau) \times \mathbf{h}_2(\Delta\tau) \times \cdots \times \mathbf{h}_N(\Delta\tau)| \quad (16)$$

For $N = 1$ to 4, an N th-order Lyapunov exponent can be defined by

$$\lim_{\substack{t \rightarrow \infty \\ \delta x \rightarrow 0}} \frac{1}{t} \log \frac{V_N(t)}{\delta x_0^N} = \lambda_N^* \quad (17)$$

To avoid divergence problems, the vectors $\mathbf{h}_i$ are orthogonalized and normalized to length δx_0 after each interval $\delta\tau$, thus generating a new set of basis vectors, and the process is repeated. At each step, the logarithm in equation (17) is evaluated and the results are time-averaged until convergence is reached, as shown in Fig. 2b. Knowing the λ_N^* , it is a simple matter to determine the Lyapunov exponents λ_i using

$$\lambda_N^* = \sum_{i=1}^N \lambda_i, \quad N = 1 \text{ to } 4 \quad (18)$$

The procedure thus bypasses the need to take derivatives and obtain eigenvalues of the jacobian.

The results are listed in Table 1. The numerical and analytic exponents λ_1 , the quantity responsible for chaos, differ by only 9%, as does λ_4 in each case. The disparity in the pair of $\lambda_{2,3}$ values is indicative of the level of accuracy for these quantities, which are an order of magnitude smaller and not important for chaos. According to the sum rule in equation (11), the sum should equal -2Γ , which they do within an uncertainty of 2%. In addition, because of pairwise symmetry, $\lambda_2 + \lambda_3 = \lambda_1 + \lambda_4 = -\Gamma$. Clearly, the agreement further supports the conclusions derived from the analytic results.

Farmer *et al.*¹⁴ have advanced a heuristic argument for a Lyapunov dimension which they define as

$$d_L = k + \sum_{i=1}^k \lambda_i / (-\lambda_{k+1}) \quad (19)$$

where k is the largest integer that satisfies $\sum_{i=1}^k \lambda_i > 0$. Using the numerical results in Table 1, the Lyapunov dimension is $d_L = 3.95$. If the system is not chaotic, we expect $d_L = 4$, corresponding

to the four degrees of freedom. Thus, not all of phase space is accessible, which is another signature of chaos.

Strange attractor

The motion in the chaotic state is very complicated in general, and to characterize it some simplifying procedure is required. One means is a return map, such as is shown in Fig. 3, where the relative two-ion position at one instant z_n is plotted against its subsequent position z_{n+1} for a fixed time interval. If the time interval is made synchronous (stroboscopic) with a periodic motion of the two ions (for example, the micro-motion), the simple patterns or so-called limit cycles of Fig. 3 appear. If the two ions were attracted to a single limit cycle, they would be in the ordered rather than the chaotic state. Chaotic behaviour arises because collisions transfer the ions from one limit cycle to another. In Fig. 3, for example, three collisions have occurred to produce three new orbits. Verification that a collision has actually taken place is evident from the trajectories in the z - r plane (non-stroboscopic), which always correlate with a jump to a new limit cycle. Thus, the distribution of collision times, which is deterministic, introduces the random element of two-ion chaos.

By extending the duration of these calculations to $\sim 2 \times 10^6$ reduced time units, the above limit cycles 'fill in' to form the galaxy-like spirals of Fig. 4, which apparently represents a new type of attractor—a two-ion strange attractor. The plot was obtained with a PS/2 computer and a routine¹⁵ modified for the two-ion equations of motion. The time evolution can be followed rapidly or point by point. The latter reveals the orbits of Fig. 3 and collision-induced jumps to new orbits. The merging of these orbits in a spiral structure is due to damping, which results in a slow contraction along spiral arms—a curious topology.

In the centre of the pattern is a sixfold symmetrical figure that depends on the magnitude of the q parameter. This seems to be associated with the phenomenon of frequency locking. As q increases, the trap's potential well gets steeper, forcing the two ions closer together, and the nonlinear Coulomb force increases, causing shifts in the frequencies of motion. Frequency locking occurs when these frequencies are in a ratio of integers, whereupon chaos gives way to order. For certain initial conditions and values of q , one finds the frequency locking shown in Fig. 5, where 19 stable points remain after transients have decayed. The degree of symmetry (the number of points) depends on the winding number w , the ratio of integers of the micro-motion to the secular frequencies; in this case, a Fourier analysis of the motion reveals that $w = 19/3$, explaining why 19 points are seen in Fig. 5. The frequency-locked attractor is not sensitive to collisions because its volume in phase space excludes

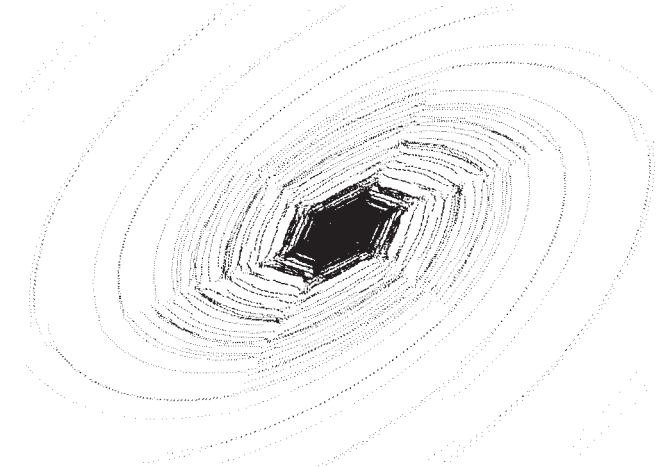


FIG. 4 The two-ion strange attractor plotted as the return map r_{n+1} against r_n strobed at the micro-motion period. Here, 1.98 million points have been sampled, $(r_0, z_0, \dot{r}_0, \dot{z}_0) = (5, 0, 2, 2)$, $q = 0.9$ and $\Gamma = 0.0005$.



FIG. 5 An example of frequency locking in a plot of r_{n+1} against r_n strobed at the micro-motion period. The winding number $w = 19/3$, the ratio of micro-motion to secular frequencies. Conditions are $(r_0, z_0, \dot{r}_0, \dot{z}_0) = (5, 0, 2, 2)$, $q = 0.86$ and $\Gamma = 0.0005$.

the region near the origin. Small changes in q can change w , destroying the ratio of integers and inducing a change from a frequency-locked to a chaotic mode. Thus, the central pattern of Fig. 4 seems to be an approximation to frequency locking which cannot be sustained with increasing distance from the centre of the pattern. Notice also that at large distances, folding occurs along these symmetry axes.

We have not yet fully explored the basins of attraction, although we find that for the same q value, different initial conditions can lead either to a strange attractor or to a frequency-

locked state. Once frequency locking occurs, it remains indefinitely because collisions cannot then occur. On the other hand, there is some uncertainty as to whether the chaotic motion is a transient that ultimately passes to a locked configuration; the only test seems to be that of longer integration times. Typically, integration times of 100,000 scaled units were sufficient to test whether a chaotic pattern was persistent (Fig. 4 has an integration time of 2×10^6 steps). Also, because of the sensitivity to initial conditions, programs with different round-off errors will produce different results. $\square$

Received 29 August 1989; accepted 5 February 1990.

1. Poincaré, H. *Les methodes nouvelles de la mecanique celeste* (Gauthier-Villars, Paris, 1892).
2. Delande, D. & Gay, J. C. in *Atomic Physics* Vol. 11 (eds Haroche, S., Gay, J. C. & Grynberg, G.) 281–297 (World Scientific, New Jersey, 1989).
3. Hoffnagle, J., DeVoe, R. G., Reyna, L. & Brewer, R. G. *Phys. Rev. Lett.* **61**, 255–258 (1988).
4. Wuerker, R. F., Shelton, H. & Langmuir, R. V. *J. appl. Phys.* **30**, 342–349 (1959).
5. Blumel, R. *et al. Nature* **334**, 309–313 (1988).
6. Berge, P., Pomeau, Y. & Vidal, C. *Order Within Chaos* 279 (Wiley, New York, 1984).
7. Lichtenberg, A. J. & Lieberman, M. A. *Regular and Stochastic Motion* 262 (Springer-Verlag, New York, 1983).

8. Dehmelt, H. G. *Adv. atom. mol. Phys.* **3**, 53–72 (1967).
9. Kapitza, P. L. *Zh. Eksperim. i Teor. Fiz.* **21**, 588 (1951).
10. Landau, L. D. & Lifshitz, E. M. *Mechanics* 93 (Pergamon, New York, 1978).
11. Abramowitz, M. & Stegun, I. A. *Handbook of Mathematical Functions* 724 (National Bureau of Standards, Washington, DC, 1968).
12. Blumel, R., Kappler, C., Quint, W. & Walther, H. *Phys. Rev. A* **40**, 808–823 (1989).
13. Shimada, I. & Nagashima, T. *Prog. theor. Phys.* **61**, 1605–1616 (1979).
14. Farmer, J. D., Ott, E. & Yorke, J. A. *Physica D* **7**, 153–180 (1983).
15. Yorke, J. A. *Dynamics, A Program for IBM PC Clones* (University of Maryland, College Park, 1988).

Developmental regulation of human fetal-to-adult globin gene switching in transgenic mice

Tariq Enver*, Natacha Raich*, Allen J. Ebens*, Thalia Papayannopoulou†, Frank Costantini‡ & George Stamatoyannopoulos*

Divisions of * Medical Genetics and † Hematology, Department of Medicine, University of Washington, Seattle, Washington 98195, USA
‡ Department of Genetics and Development, College of Physicians and Surgeons, Columbia University, New York, New York 10032, USA

Transgenic mice containing a human fetal (γ -) or adult (β -) globin gene linked to the β -globin gene locus activation region (LAR) express the gene throughout development. By contrast, transgenic mice containing LAR linked to both a fetal and an adult globin gene display the normal developmental switch from fetal to adult gene expression. This suggests that the human fetal-to-adult globin gene switch is controlled through a mutually exclusive interaction between LAR and either the γ - or β -globin gene, resulting in the expression of only one gene at any given moment.

DIFFERENT members of the human β -globin gene cluster are expressed in embryonic, fetal and adult erythroid cells. This phenomenon is termed haemoglobin switching (see ref. 1 for a review). The β -globin gene locus comprises five functional genes arranged (5'–3') (ϵ , γ , α , δ , β). A series of erythroid-specific superhypersensitive sites are located 6–20 kilobases (kb) upstream of the ϵ -globin gene^{2–4}. This region, termed locus activation region (LAR)⁵ or dominant control region (DCR)⁶, is thought to be largely responsible for activating the β -globin gene domain and facilitating high-level globin gene transcription in erythroid cells^{5–12}. Linkage of the β -globin gene LAR to heterologous erythroid genes (α -globin gene)¹³ or housekeeping genes (thymidine kinase gene)^{7,9} results in their high-level expression. Most strikingly, LAR can reprogramme the tissue specificity of expression of nonerythroid lineage-specific genes⁷.

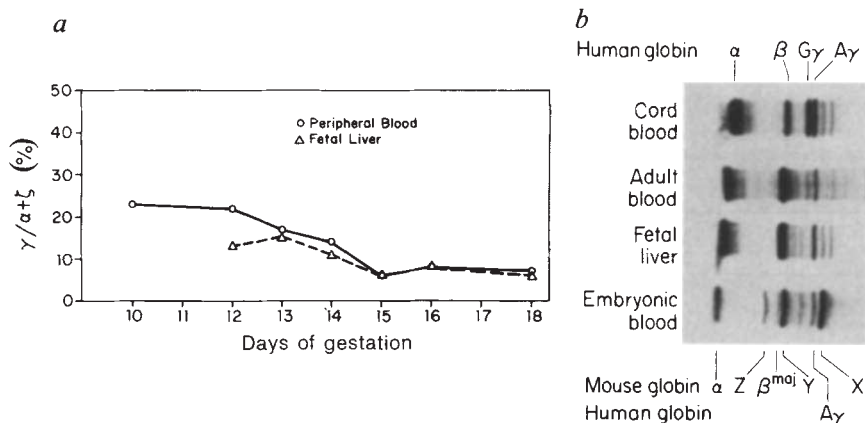
For example, linkage to the β -globin-gene LAR confers an erythroid pattern of expression on the lymphoid-specific gene *Thy-1*⁷. The dominant control capability of LAR extends to the regulation of developmental stage-specific gene expression¹⁴. In the absence of LAR, the human γ -globin gene has an exclusively embryonic pattern of expression in transgenic mice^{15,16}. But when linked to LAR it is expressed in yolk sac-derived embryonic erythrocytes, as well as in fetal-liver and adult bone-marrow-derived definitive erythroid cells¹⁴. Given this dominant nature of LAR, appropriate developmental regulation of fetal and adult β -like globin genes necessitates that they escape the influence of LAR at the ontogenetic stages in which they are quiescent. We have now produced transgenic mice containing either the human γ - or β -globin gene linked to a LAR cassette, and show that neither gene can evade the LAR effect and that they are expressed at similar levels throughout development. But we find that developmental regulation is restored if the γ - and β -globin genes are linked to LAR as part of a 29-kb fragment that preserves their genomic structure. These results suggest a reciprocal mechanism for globin gene switching whereby the fetal and adult globin genes compete for the influence of LAR.

LAR constructs that defy developmental control

Transgenic mice containing the μ LAR γ construct, in which the γ -globin gene is linked to a 2.5-kb cassette¹⁰ containing four of the erythroid-specific hypersensitive sites from LAR, express the γ -globin gene in both primitive and definitive erythroid cells¹⁴. A comparative analysis of the level of accumulated γ -globin chains in day-14 blood (contains persisting yolk sac-derived embryonic erythroid cells) and fetal liver (site of definitive erythropoiesis) of the same transgenic embryos indicated that the level of γ -globin gene expression is similar at

FIG. 1 Developmental γ -globin gene expression in μ LAR $^{\gamma}$ transgenic mice. *a*, Graphs of human γ -globin-chain expression as a function of total mouse α -globins in developmentally staged μ LAR $^{\gamma}$ transgenic embryos. The data were derived from densitometric scanning of isoelectric focusing gels of ^3H -labelled globin chains from the peripheral blood and fetal liver. *b*, Examples of the data from day-10 peripheral embryonic blood, day-14 fetal liver and adult peripheral blood and human cord blood to indicate the migration of the human $^{\gamma}$ -globin chain. Murine globin chains X, Y and Z derive from the ζ , ϵ^{γ} and βH1 genes, respectively.

METHODS. Adult transgenic mice containing the μ LAR $^{\gamma}$ construct were made as previously described²⁸. This construct is described in detail in ref. 14. Single cell suspensions from the blood and livers of developmentally staged transgenic offspring of one such adult male were incubated with 75 μCi [^3H]leucine; cells were then lysed, and the radiolabelled globin chains analysed by isoelectric focusing. After autoradiography gels were fixed, dried and fluoro-



graphed¹⁴. Quantitation of γ -globin chain expression was achieved by densitometric scanning of fluorograms.

both of these developmental stages¹⁴. To examine γ -globin expression during murine ontogeny in more detail, we established lines of mice containing the μ LAR $^{\gamma}$ construct and examined γ -globin gene expression in staged embryos obtained from their breeding. Establishing lines of mice containing LAR/globin gene recombinants is difficult^{6,17}, because the high level of human globin in murine erythroid cells results in a lethal chain imbalance¹⁷. We took advantage of the fact that the 2.5-kb μ LAR cassette previously used is not as efficient at directing high-level expression as the native LAR¹⁴. Therefore although position-independent copy number-dependent expression is observed, the absolute expression per gene is sufficiently moderate so as not to jeopardize the viability of animals containing low-to-medium copy numbers of the transgene. We obtained seven μ LAR $^{\gamma}$ transgenic adults, and mated one containing three intact copies of the transgene with several females. These females were killed at recorded times after the detection of vaginal plugs, and their fetuses removed. Erythroid material (peripheral blood or fetal liver, or both) was isolated from the transgenic embryos and incubated in the presence of tritiated leucine to label the globin chains, which were then analysed by isoelectric focusing. In Fig. 1*a* the quantity of γ -globin chain at various stages of mouse development is expressed as a fraction of total mouse α -globin ($\gamma/(\alpha+\zeta)$). Examples of the isoelectric focusing analyses from which these data were derived are shown in Fig. 1*b*. These results show that, although there is a slight modulation (about threefold) of γ -globin expression during murine ontogeny, the μ LAR $^{\gamma}$ construct is not tightly developmentally regulated, and is expressed at similar levels in yolk sac-derived embryonic erythroid cells, as well as fetal-liver and bone-marrow-derived definitive erythroid cells.

To investigate whether LAR could override the developmental control of the human β -globin gene we used the μ LAR $^{\beta}$ construct, in which the 2.5-kb μ LAR cassette is linked to a 4.8-kb fragment containing the β -globin gene. This fragment is developmentally regulated in transgenic mice: it is expressed only in definitive and not in primitive erythroid cells derived from the yolk sac^{16,18-22}. We studied the expression of the β -globin transgene at embryonic day 11, fetal day 14, and in adult mice. At day 11, the peripheral blood of the mouse embryo contains exclusively primitive erythroid cells derived from the yolk sac of the embryo. At day 14, the fetal liver is the primary site of definitive erythropoiesis and is largely composed of definitive erythroblasts. The peripheral blood of adult mice is composed of definitive erythroid cells derived from the bone marrow (see refs 15 and 18). From several independent injections we obtained six day-11 transgenic embryos, three transgenic fetuses at day 14, and four transgenic adults. We prepared cellular smears of erythroid cells from these different developmental stages and stained them with fluorescently labelled monoclonal antibodies specific for the human β -globin chain. As observed in previous experiments without LAR²², definitive erythroid cells of both adult blood and fetal liver express human β -globin (data not shown). In marked contrast to these earlier experiments²², adult human globin was also apparent in the nucleated embryonic erythroid cells of the mouse (Fig. 2). We confirmed these results at the RNA level by RNase protection of day-11 and adult peripheral blood RNAs, using probes specific for human β - and mouse α - and ζ -globin gene messenger RNAs. Figure 3 shows the results of these analyses in similar-copy-number μ LAR $^{\beta}$ transgenic mice at different developmental stages. These data show that in the presence of LAR, the β -globin gene loses its

FIG. 2 Expression of μ LAR $^{\beta}$ in embryonic erythroid cells. Cellular smears of placental blood from day-11 μ LAR $^{\beta}$ transgenic embryos labelled with a fluorescently conjugated monoclonal antibody specific for human β -globin chains. The placental blood used for staining contained embryonic erythroblasts and was contaminated with many maternal cells. Note that the large embryonic erythroid cells of the transgenic embryo are brightly stained, whereas the small enucleated adult erythrocytes of the nontransgenic mother are not stained.

METHODS. The μ LAR $^{\beta}$ construct contains the 2.5-kb μ LAR cassette linked to a 4.8-kb *Bgl*II-*Eco*RV fragment containing a marked copy of the human β -globin gene so that the genomic orientation of the two elements is preserved. Preparation and antibody staining of cellular smears was performed as previously described¹⁴.

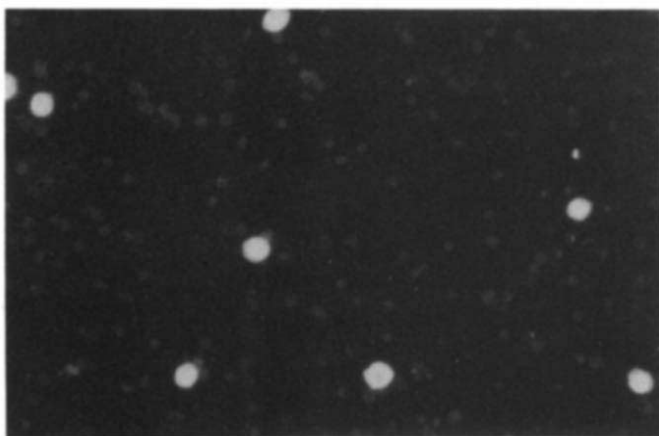
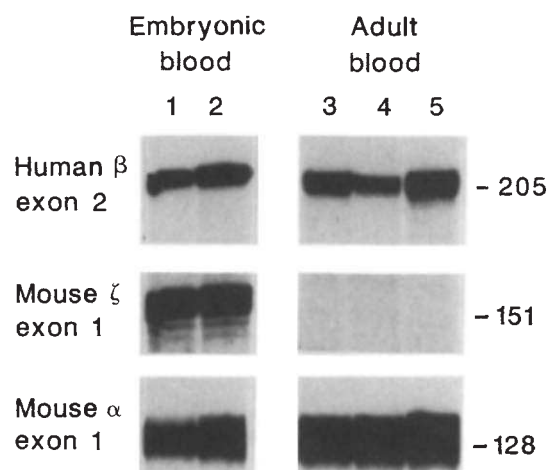


FIG. 3 RNase protection analysis of β -globin gene expression in μ LAR β transgenic animals, using probes specific for human β -, mouse ζ - and mouse α -globin mRNAs. Lanes 1 and 2 protection analyses of peripheral blood RNA from two different day-11 μ LAR β transgenic embryos. Lanes 3, 4 and 5, protections analyses of peripheral blood RNA from three adult transgenic animals. The probes used and sizes of the protection products in base pairs are indicated (right).

METHODS. RNA was prepared from peripheral blood as described²⁹. Portions (200 ng) of RNA were used for each protection analysis, performed as previously described³⁰. Probes used: pM α (ref. 31), pM ζ (ref. 31) and pT7 β^m (T.E., unpublished).



developmental regulation, because it is expressed at similar levels in all stages of murine ontogeny.

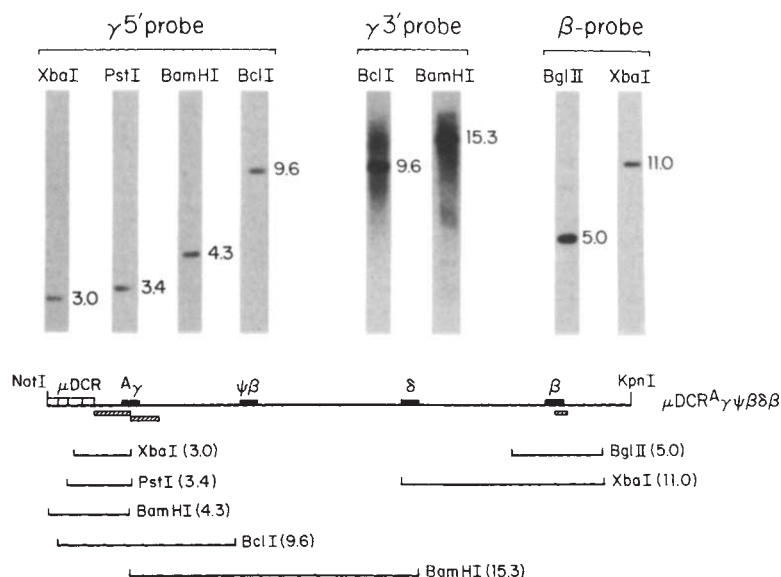
LAR constructs restoring developmental control

Our experiments with μ LAR γ and μ LAR β constructs raise a paradox, because both the γ - and β -globin genes are developmentally regulated when in their normal chromosomal context, which includes LAR. To mimic the genomic organization of fetal-adult globin genes in the globin locus we extended the boundary of the μ LAR γ construct (Fig. 4) so that 29 kb of contiguous genomic globin locus sequences were contained up to and including the β -globin gene. We used the purified eukaryotic insert from this construct (μ LAR $\gamma\psi\beta\delta\beta$) to produce day-11, day-14 and adult transgenic mice. Transgenic animals were identified and analysed for fidelity of transgene integration by Southern blotting (an example is shown in Fig. 4). The expression of the human γ - and β -globin genes was documented at the cellular level by staining with monoclonal antibodies specific for human γ - and β -globin chains (anti- γ and anti- β antibodies, respectively; Fig. 5). The blood of the day-11 embryo is exclusively composed of embryonic erythroblasts derived from the yolk sac. These cells fluoresced intensely after staining with anti- γ antibody (Fig. 5a), whereas only an occasional embryonic erythroblast was labelled with anti- β monoclonal antibody (Fig.

5b). Essentially the opposite results were obtained when the blood of adult μ LAR $\gamma\psi\beta\delta\beta$ transgenic animals was analysed. In this case, most of the adult erythroid cells were labelled with the anti- β antibody (Fig. 5f), whereas only a few cells were labelled with the anti- γ antibody (Fig. 5e). But in the liver of day-14 fetus, the number cells containing γ -globin chains was similar to that containing β -globin chains (Fig. 5c, d), resembling a situation similar to the perinatal period of globin gene switching in man. We confirmed these results at the molecular level by RNase protection of RNA derived from embryonic (day 11), fetal (day 14) and adult erythroid cells, using probes specific for the human γ - and β -globin mRNAs (Fig. 6a). Because the γ - and β -globin genes are linked on the same construct, we could derive the ratio of fetal to adult globin mRNA for each individual animal and use these for comparisons. A graph of these results is shown in Fig. 6b. These data, like those obtained by antibody staining, show that in embryonic blood cells the γ -globin gene is predominantly expressed; in the fetal liver, the γ - and β -globin genes are expressed at similar levels, whereas in the peripheral blood of the adult mouse, expression of the β -globin gene predominates. There is a 100-fold change in the relative expression of the γ - and β -globin genes between the embryonic and adult stages of erythropoiesis in the mouse (Fig. 6b). Taken together, these results show that correct

FIG. 4 Structural analysis of the μ LAR $\gamma\psi\beta\delta\beta$ transgenic mice by Southern blotting. Top, genomic DNA prepared from the carcasses of transgenic animals was digested with the restriction enzymes indicated and Southern blotted using the probes from 5' and 3' portions of the human γ -globin gene (γ 5' and γ 3' probes), and the 3' portion of the human β -globin gene (β -probe). Bottom, location of the restriction sites and probes used, together with the sizes (kb) of the expected fragments.

METHODS. Construct μ LAR $\gamma\psi\beta\delta\beta$ was built as follows. The μ LAR γ fragment containing the 2.5-kb μ LAR cassette linked to the 3.3-kb *HindIII* fragment containing the γ -globin gene was transferred as a *NotI*-*SaI* fragment into a version of cosmid pHC79 (see ref. 32) in which a *NotI* site was created at the *EcoRI* position. This construct pHC79 μ LAR γ was linearized at the *XhoI* site in the first exon of the γ -globin gene. The cosmid cCH1 (gift of Richard Gelinas) contains a 39-kb *KpnI* human globin locus fragment spanning 5' of γ to 3' of β , in a modified pHC79 vector in which the *BamHI* site was converted to a *KpnI* site. Cosmid cCH1 was also cleaved with *XhoI* and a fragment containing the γ -to- β region together with the vector was isolated and ligated to *XhoI*-cleaved pHC79 μ LAR γ . The packaged ligation mix was used to transform bacteria. Bacteria containing the appropriate recombinant molecule were initially identified by colony screening³³ using probes specific for LAR and β -globin, and confirmed by extensive restriction mapping of cosmid miniprep DNA. Southern blotting and preparation of genomic DNA was performed as described^{14,33}. Probes used: γ 5'(1.8-kb *HindIII*-*BamHI* γ -fragment), γ 3'(1.4-kb *BamHI*-*HindIII* γ -



fragment), β (0.7-kb *EcoRI*-*PstI* fragment from the 3' region of the β -globin gene), and were labelled as described¹⁴. Final blot washes: 0.1 $\times$ SSC for 30 min at 65 $^{\circ}$ C.

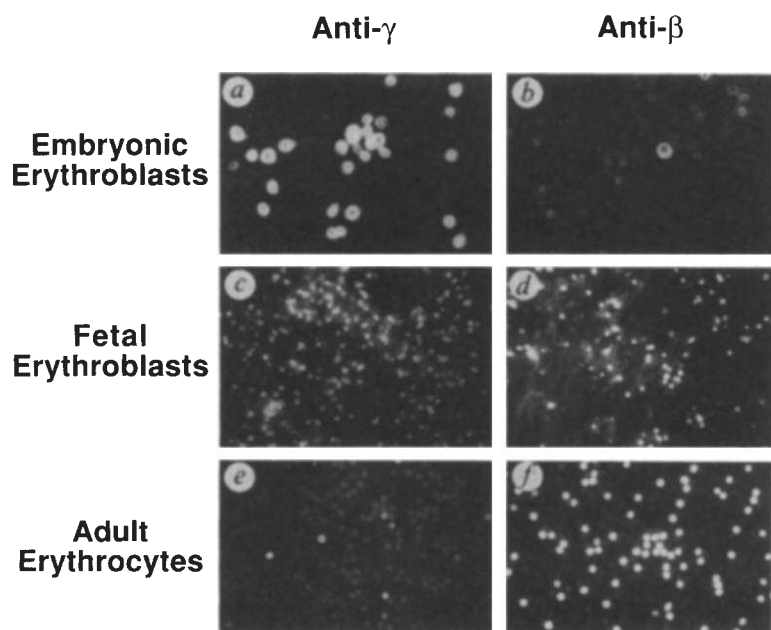


FIG. 5 Cellular basis of the fetal-to-adult globin gene switch in $\mu\text{LAR}^{\Delta\gamma\psi\beta\delta\beta}$ transgenic mice. Cellular smears were prepared from day-11 peripheral blood (a, b), day-14 fetal liver (c, d) and adult peripheral blood (e, f) of transgenic animals. These smears were stained with fluorescently conjugated anti- γ (a, c, e) or anti- β (b, d, f) monoclonal antibodies.

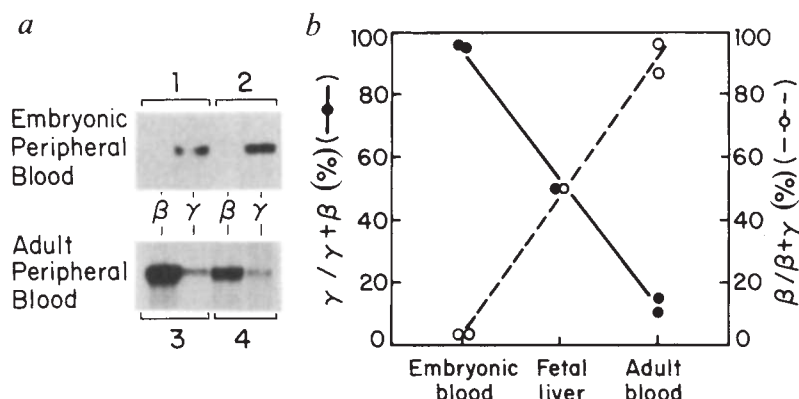


FIG. 6 Analysis of $\mu\text{LAR}^{\Delta\gamma\psi\beta\delta\beta}$ expression by RNase protection. a, Portions (200 ng) of RNA prepared from the peripheral blood of day-11 transgenic embryos (mice 1 and 2) or adults (mice 3 and 4) were analysed with probes specific for exon2 of either the human γ - or β -globin gene mRNAs as indicated. b, Graph of γ - or β -globin gene mRNA expression plotted as a fraction of total human non- α globin mRNA during mouse ontogeny.

METHODS. RNase protections were performed as described in Fig. 3 legend. Values for γ - and β -globin gene expression were obtained by densitometric scanning as normalized by comparison with mouse α - and ζ -globin gene expression in the same samples, also determined by RNase protection.

globin gene switching is achieved when both the γ - and β -globin genes are present on the same construct, together with LAR.

Implications for globin gene switching

Previous studies have shown that when fetal and adult human globin genes are introduced individually into transgenic mice they are developmentally regulated. This implies that stage-specific transacting factors interact with the genes themselves to bring about globin gene switching. We have shown that linkage of LAR to individual globin genes invalidates these interactions and results in constitutive expression of the genes throughout ontogeny. LAR presumably achieves this either by overriding the effect of stage-specific repressors or by alleviating the need for stage-specific activators. LAR is unable to do this when both genes are linked together on a 29-kb genomic globin locus fragment. In this case, developmental regulation of the genes is restored. Presumably the sequences present in this fragment prevent the productive interaction between LAR and the β -globin gene in primitive erythroid cells, and between LAR and the γ -globin gene in adult erythroid cells. The nature of these sequences and the mechanism by which they negate LAR/globin gene interactions is unknown. We speculate that the human fetal-to-adult globin gene switch may be effected in a manner similar to the embryonic-to-adult globin gene switch in chickens^{23,24}. In the chicken system, switching is brought about by the competition of two promoters for a common enhancer. The outcome of this contest is thought to depend on the availability of developmentally regulated promoter-specific

transacting factors that alter the ability of the promoters to use the enhancer. Human fetal-to-adult globin gene switching would then depend on stage-specific transacting factors interacting with the globin genes and changing their relative abilities to engage LAR. As in the chicken system, a key component of this model is that LAR interacts with only one gene at a time. This is consistent with mechanisms of LAR action based on the juxtaposition of LAR and the target gene by DNA looping, or the absorbance of LAR-generated torsional stress by DNA melting²⁵.

A competitive model of human globin gene switching can also explain a group of human globin-gene developmental mutants that are collectively known as hereditary persistence of fetal haemoglobin (HPFH, mutants (see ref. 1 for a review). Individuals with these mutations continue to express a high level of fetal haemoglobin in adult life. The mutations underlying this phenotype are diverse at the molecular level but fall into two main categories. In one group (deletion HPFH mutations) there are large deletions in the 3' region of the β -globin gene locus that remove the adult globin genes (δ - and β -globin genes) but leave the fetal globin genes (γ - and ζ -globin genes) intact. By contrast, the nondeletion HPFH mutations preserve the structural integrity of the β -globin gene locus. However, this group has point mutations in the fetal globin genes that could alter both the factor-binding and transcriptional efficiency²⁶ of the affected promoters. With deletion HPFH mutants, in absence of competing adult promoters, the action of LAR could be engaged by the remaining fetal globin genes resulting in their expression in adult life. In nondeletion HPFH mutants, we

propose that the point mutations alter the nature of the transcriptional complex so as to increase the ability of the fetal globin promoter to compete effectively with the adult β -globin gene

promoter for the action of LAR. This notion is supported by the observation that the β -globin gene *in cis* configuration with the HPFH mutant γ -gene is hypofunctional²⁷. □

Received 8 January; accepted 2 February 1990.

1. Stamatoyannopoulos, G. & Nienhuis, A. W. in *Molecular Basis of Blood Diseases* (eds Stamatoyannopoulos, G., Nienhuis, A. W., Leder, P. & Majerus, P. W.) 66–93 (Saunders, Philadelphia, 1987).
2. Li, Q., Powers, P. A. & Smithies, O. *J. Biol. Chem.* **260**, 14901–14910 (1985).
3. Tuan, D., Solomon, W., Li, Q. & London, I. M. *Proc. natn. Acad. Sci. U.S.A.* **82**, 6384–6388 (1985).
4. Forrester, W. C., Thompson, C., Elder, J. T. & Groudine, M. *Proc. natn. Acad. Sci. U.S.A.* **83**, 1359–1363 (1986).
5. Forrester, W. C., Takegawa, S., Papayannopoulou, Th., Stamatoyannopoulos, G. & Groudine, M. *Nucleic Acids Res.* **15**, 10159–10177 (1987).
6. Grosfeld, F., van Assendelft, G. B., Greaves, D. R. & Kollias, G. *Cell* **51**, 975–985 (1987).
7. Van Assendelft, G. B., Hanscombe, O., Grosfeld, F. & Greaves, D. R. *Cell* **56**, 969–977 (1989).
8. Ryan, T. M. *et al. Genes Dev.* **3**, 319–323 (1989).
9. Talbot, D. *et al. Nature* **338**, 352–355 (1989).
10. Forrester, W. C., Novak, U., Gelinas, R. & Groudine, M. *Proc. natn. Acad. Sci. U.S.A.* **86**, 5439–5443 (1989).
11. Tuan, D. Y. H., Solomon, W. B., London, I. M. & Lee, D. P. *Proc. natn. Acad. Sci. U.S.A.* **86**, 2554–2558 (1989).
12. Curtin, P. T., Liu, D., Liu, W. Chang, J. C. & Kan, Y. W. *Proc. natn. Acad. Sci. U.S.A.* **86**, 7083–7086 (1989).
13. Ryan, T. M., Behringer, R. R., Townes, T. M., Palmiter, R. D. & Brinster, R. L. *Proc. natn. Acad. Sci. U.S.A.* **86**, 37–41 (1989).
14. Enver, T., Ebens, A. J., Forrester, W. C. & Stamatoyannopoulos, G. *Proc. natn. Acad. Sci. U.S.A.* **86**, 7033–7037 (1989).
15. Chada, K., Magram, J. & Costantini, F. *Nature* **319**, 685–689 (1986).
16. Kollias, G., Wrighton, N., Hurst, J. & Grosfeld, F. *Cell* **46**, 89–94 (1986).
17. Hanscombe O. *et al. Genes Dev.* **3**, 1572–1581 (1989).

18. Magram, J., Chada K. & Costantini, F. *Nature* **315**, 338–340 (1985).
19. Townes, T. M., Lingrel, J. B., Chen, H. Y., Brinster, R. L. & Palmiter, R. D. *EMBO J.* **4**, 1715–1723 (1985).
20. Kollias, G., Hurst, J., deBoer, E. & Grosfeld, F. *Nucleic Acids Res.* **15**, 5739–5747 (1987).
21. Behringer, R. R., Hammer, R. E., Brinster, R. L., Palmiter, R. D. & Townes, T. M. *Proc. natn. Acad. Sci. U.S.A.* **84**, 7056–7060 (1987).
22. Costantini, F. *et al. Cold Spring Harb. Symp. quant. Biol.* **50**, 361–370 (1985).
23. Choi, O. R. & Engel, J. D. *Cell* **55**, 17–26 (1988).
24. Nickol, J. M. & Felsenfeld, G. *Proc. natn. Acad. Sci. U.S.A.* **85**, 2548–2552 (1988).
25. Patient, R. K. & Allan, J. *Curr. Opin. Cell Biol.* **1**, 454–459 (1989).
26. Rixon, M. & Gelinas, R. *Molec. cell. Biol.* **8**, 713–721 (1988).
27. Papayannopoulou, Th., Enver, T., Takegawa, S., Anagnou, N. P. & Stamatoyannopoulos, G. *Science* **242**, 1056–1058 (1988).
28. Hogan, B. L. M., Costantini, F. & Lacy, E. *Manipulating the Mouse Embryo: A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1986).
29. Karlinsey, J., Stamatoyannopoulos, G. & Enver, T. *Analyt. Biochem.* **180**, 303–306 (1989).
30. Enver, T. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 9091–9095 (1988).
31. Baron, M. H. & Maniatis, T. *Cell* **46**, 591–602.
32. Gelinas, R. E., Endlich, B., Pfeiffer, C., Yagi, M. & Stamatoyannopoulos, G. *Nature* **313**, 323–325 (1985).
33. Maniatis, T., Fritsch, E. F. & Sambrook, J. in *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).

ACKNOWLEDGEMENTS. The authors are grateful to Betty Josephson, Betty Nakamoto, Chu-Hui Peng, and Karin Gale for outstanding technical assistance. Our thanks are also due to Joan Carpenter and Kathy Forbush of the University of Washington transgenic mouse facility under the direction of Roger Perlmutter for producing several of the mice used in this study. This work was supported by NIH grants and a Cooleys Anemia Foundation award to T.E.

Recurrence times of γ -ray bursts from neutron star/comet encounters

I. G. Mitrofanov & R. Z. Sagdeev

Space Research Institute, Profsojuznaja 84/32, Moscow 117810, USSR

RECENT observations¹ from Phobos 2 added about one hundred γ -ray bursts to the several hundred detected over the past 20 years. One of the most promising explanations of these bursts² is that they originate from the activation of the magnetospheres of old neutron stars³, for example by encounters with small bodies. Here we use estimates of the number density of small bodies, either interstellar or in circumstellar Oort-like clouds, to predict the recurrence times of γ -ray bursts from individual sources, on the assumption that they are produced from neutron star/comet encounters (direct impacts^{4,5}, captures^{6,7} or passages⁸). We find that 'classical' (non-repeating) bursts may be accounted for by passages of interstellar small bodies through the magnetosphere, whereas the small number of repeating bursts is consistent with the interaction of neutron stars with small bodies within Oort-like clouds. In the latter case, the best agreement with observational statistics is found by adopting the recent model of the Oort cloud⁹ derived from the study of comet Halley by the Vega spacecraft¹⁰.

The idea that γ -ray bursts arise from the interactions of neutron stars and small bodies from interstellar space^{4–8} has the advantage that it does not require permanent accretion or spontaneous cracking to explain the sporadic activity of neutron stars. It was proposed that the energy of γ -ray bursts (GRBs), $E_{\text{GRB}} > 10^{36}$ erg, was supplied either by direct collisions^{4,5} or by the capture of small bodies by strong tidal or magnetic drags^{6,7}. The assumed number density of interstellar small bodies is, however, too small to account for either the observed rate of 'classical' bursts¹¹ or for the short recurrence time of repeating bursts. If the neutron star retains the pre-supernova Oort-like cometary cloud¹², the frequency of the collisions would be increased, but this cannot be used to explain the number of non-repeating bursts, which are the most common type.

According to a model of pulsar family evolution³, a 'short circuit' inside the outer gaps of a dead pulsar of period > 200 ms might result in the emission of a γ -ray burst. The passage of a small body through the outer regions of the magnetosphere may lead to such a short circuit⁸. Here we estimate recurrence times for three types of encounters between neutron stars and small bodies. In the absence of any knowledge about the number-mass distribution of interstellar small bodies, we have used two models^{9,13} of comets in the Oort cloud of the Solar System.

In the first model¹³, which uses the previously accepted value of 0.6 for the average albedo A of comets, the total number of small bodies with masses greater than M in the Oort cloud is

$$N_{\text{sb}}^{\text{Oort}}(> M) \approx 5 \times 10^{10} (M/M_*)^{1-\alpha} \quad (1)$$

where $\alpha = \alpha_1 = 2.48$ and $\alpha = \alpha_2 = 1.73$ correspond to $M > M_*$ and $M < M_* = 4 \times 10^{16}$ g, respectively (here M_* is the average cometary mass). The only directly measured comet albedo, that of comet Halley, is $A' = 0.04$ (ref. 10). Assuming comet Halley to be representative, a new model for the Oort cloud¹⁰ has been proposed in which the masses of comets are $(A/A')^{3/2} \approx 60$ times larger. For this model one can also use equation (1), but with $M_* \approx 2 \times 10^{18}$ g.

The density of interstellar small bodies Σ_{sb} cannot exceed the upper limit of $\Sigma_{\text{sb}}^{\text{max}} = 3 \times 10^{-3} M_{\odot} \text{ pc}^{-3}$, which is based on the ratio of interstellar hydrogen and heavy elements¹¹. Assuming this maximum density, we estimate the number density of small bodies to be

$$n_{\text{sb}}(> M) = \frac{\Sigma_{\text{sb}}^{\text{max}} (2 - \alpha_2) (\alpha_1 - 2)}{(\alpha_1 - \alpha_2) (\alpha_2 - 1) M_*^{2-\alpha_2} M^{\alpha_2-1}} \quad (2)$$

for both models.

The mean time between two encounters of a neutron star with a small body, $t_{\text{rec}} = (\pi R^2 V n_{\text{sb}}(> M))^{-1}$, depends on the impact parameter R and their relative velocity V at infinity. In Table 1 we present estimates of t_{rec} for both models and assuming three kinds of interactions: (1) direct collisions of neutron stars with small bodies^{4,5}; (2) magnetic or tidal captures of small bodies^{6,7} (both cases with $M > 10^{16}$ g); and (3) passages of small bodies through the light-cylinder region⁸ of the neutron-star magnetosphere ($M > 10^{12}$ g and $M > 10^{16}$ g; see below).

In almost all cases bursts from particular sources have been

TABLE 1 Mean recurrence time of γ -ray bursts

Types of interaction	Interstellar space; n_{sb} from equation (2): $\sim 3 \times 10^{16} (10^{12} \text{ g}/M)^{0.73} (2 \times 10^{18} \text{ g}/M_*)^{0.27}$		Oort-like cloud; n_{sb} based on equation (1): $\sim 10^{14} (M_*/M)^{0.73}$	
	ref. 13 model $M_* = 4 \times 10^{16} \text{ g}$	ref. 9 model $M_* = 2 \times 10^{18} \text{ g}$	ref. 13 model $M_* = 4 \times 10^{16} \text{ g}$	ref. 9 model $M_* = 2 \times 10^{18} \text{ g}$
Direct impacts ^{4,5} : $R_{\text{min}} = 10^6 \text{ cm}$ and $M > 10^{16} \text{ g}$	10^7	3×10^7	3×10^6	2×10^5
Captures ^{6,7} : $R_{\text{min}} = 3 \times 10^7 \text{ cm}$ and $M > 10^{16} \text{ g}$	3×10^5	7×10^5	10^5	5×10^3
Passages (ref. 8 and this work): $R_{\text{min}} = 10^9 \text{ cm}$ and $M > 10^{12} \text{ g}$ (in brackets, $M > 10^{16} \text{ g}$)	10 (10^4)	30 (3×10^4)	4 (3×10^3)	0.2 (200)

$t_{\text{rec}} = 10^{27} R_{\text{min}}^{-1} n_{\text{sb}}^{-1} (10/V) \text{ yr}$ (where R is in cm, n_{sb} in pc^{-3} and V in km s^{-1}) for different types of interactions between neutron stars and small bodies⁴⁻⁸ and different models^{9,13} of the Oort cloud.

detected only once so the recurrence times for 'classical' γ -ray bursts are $\gg 20 \text{ yr}$. Only for the case of small bodies passing through the light-cylinder region is the frequency of encounters in interstellar space consistent with the minimum recurrence time of 'classical' bursts. Moreover, recurrence times much longer than 20 years are also quite probable and such encounters could account for the observed burst rate even for $\Sigma_{\text{sb}} \ll \Sigma_{\text{sb}}^{\text{max}}$.

During a year, about 100 intense bursts (fluences $S_0 > 10^{-5} \text{ erg cm}^{-2}$) are detected¹⁴. In our region of the galaxy the number density of burst sources is $n_s = n_{\text{ns}} \delta$, where $n_{\text{ns}} = 4 \times 10^{-3} \text{ pc}^{-3}$ is the number density of neutron stars¹⁵ and $\delta < 1$ is the fraction of them that can generate γ -ray bursts. During a year about $N_{\text{GRB}} = 100$ encounters occur in a sphere of radius

$$R_0 = \left[\frac{3 N_{\text{GRB}} t_{\text{rec}}}{4 \pi n_s} \right]^{1/3} \approx 60 \delta^{-1/3} (t_{\text{rec}}/30 \text{ yr})^{1/3} \text{ pc} \quad (3)$$

Assuming a 'standard candle' model for burst sources we associate events with fluence S_0 with objects at distance R_0 . In this case an energy of about

$$\mathcal{E}_{\text{GRB}} = 4 \pi R_0^2 S_0 \approx 4 \times 10^{36} \delta^{-2/3} \text{ erg} \quad (4)$$

is emitted in one burst, provided that the γ -rays are radiated isotropically. Events with the minimal detectable fluence, $\sim 10^{-6} \text{ erg cm}^{-2}$, correspond to a 'standard' source at a distance of $R_{\text{max}} = 3 R_0$.

For $R_0 \geq 1 \text{ kpc}$, either the quadrupole anisotropy in the galactic plane or the dipole anisotropy in the Centre-Anticentre direction should be manifested in the distribution on the sky of localized sources of bursts¹⁴. According to equation (3), such values of R_0 correspond to mean recurrence times $\geq 10^5 \text{ yr}$. Both direct impacts and captures correspond to longer recurrence times (see Table 1), and therefore to $R_0 \gg 1 \text{ kpc}$, and should result in a pronounced anisotropy of γ -ray bursts on the sky. But the isotropy of such bursts is well known¹⁴, and so we discount the importance of these kinds of interactions. The passage of a small body through the magnetosphere of a neutron star seems to be the only kind of encounter compatible with the observations. For such encounters the combination of equation (4) and the condition $R_0 < 1 \text{ kpc}$ leads to a minimum acceptable density of small bodies in interstellar space that is quite small, $\Sigma_{\text{sb}}^{\text{min}} \approx 10^{-4} \Sigma_{\text{sb}}^{\text{max}} \approx 3 \times 10^{-7} M_{\odot} \text{ pc}^{-3}$.

Among all the detected γ -ray bursts there are only three known cases of repeated events¹⁶⁻¹⁸. We cannot discount the possibility that the nature of these bursts is quite different from those of 'classical' events (see, for example, ref. 19), but we propose an alternative idea—that repeating bursts are also connected with the passage of a small body through the magnetosphere of a neutron star, but that the small body comes from the circumstellar cometary cloud.

Oort-like circumstellar clouds may be observable as infrared excesses in stellar spectra (see, for example, ref. 9) of $\sim 18\%$ of stars. Assuming that the number density of main-sequence stars

in the vicinity of the Sun is $\sim 0.13 M_{\odot} \text{ pc}^{-3}$ (ref. 20), the number density of circumstellar Oort-like clouds can be estimated to be $n_{\text{Oort}} = 2 \times 10^{-2} \text{ pc}^{-3}$. In a volume with radius $R_{\text{max}} = 3 R_0 = 200 \text{ pc}$, the total number of burst sources, which are in Oort-like clouds with mean dimensions $R_{\text{Oort}} = 10^4 \text{ AU}^{10,12}$, is about

$$2(n_{\text{Oort}}/2 \times 10^{-2} \text{ pc}^{-3})(n_{\text{ns}}/4 \times 10^{-3} \text{ pc}^{-3}) \times (R_{\text{Oort}}/10^4 \text{ AU})^3 (R_{\text{max}}/200 \text{ pc})^2 \quad (5)$$

The number density of small bodies in Oort-like clouds is much higher than that in interstellar space. Using the models in refs 9 and 13 we obtain the average periods between interactions of neutron stars with small bodies given in Table 1. Using the newer model⁹ for the cometary cloud and assuming the mechanism in which small bodies pass through the magnetospheres of neutron stars, the estimate of t_{rec} is in reasonable agreement with the average repetition time for GBS0526–66 (ref. 17) and SGR1896–20 (ref. 18).

When a comet crosses the outer gap of a neutron star, autoemission occurs from the surface due to a strong electric field E of about $10^7 B_{13} (200 \text{ ms}/P) \text{ V cm}^{-1}$ (ref. 21) (where P is the period and B_{13} is the magnetic field in units of 10^{13} G) that results in intense Joule heating and explosive evaporation. A 'short circuit' of the outer gap would be ensured by $\leq 6 \times 10^{12} E_7 (P/200 \text{ ms})$ charged particles, where E_7 is the electric field in units of 10^7 V cm^{-1} . Even a comet with mass $10^{11} \text{--} 10^{12} \text{ g}$ can provide this number if $> 10^{22}$ particles per gram are evaporated. The 'short circuit' should be balanced by large-scale currents, and the total electrostatic energy of the charge-depleted magnetosphere,

$$\mathcal{E}_{\text{el}} \approx B^2 R_{\text{ns}}^5 / (c^2 P^2) \approx 3 \times 10^{36} B_{13}^2 (200 \text{ ms}/P)^2 \text{ erg} \quad (6)$$

would be released in electromagnetic cascades resulting in a γ -ray burst.

If the proposed identification of repeating γ -ray bursts with sporadic passages of small bodies from Oort-like clouds through the magnetospheres of neutron stars is confirmed, this phenomenon could be used for studying old quiet neutron stars as well as circumstellar cometary clouds. First, a neutron star crossing a cloud could be considered as a probe for studying the number density distribution of small bodies. For example, the period of November 1983, when SGR1806–20 showed maximum activity, with a recurrence time of several days or hours (see, for example, ref. 18), could be associated with crossing of a cloud core. Second, small bodies could be considered as probes for studying the magnetospheres of old neutron stars, which usually are quiet and unobserved. And third, physical and chemical properties of small bodies could also be explored, because inside the magnetosphere a parent small body could be destroyed by non-homogeneous heating, by tidal deformations, or by magnetic forces. Intervals of very fast repetitions of γ -ray bursts from SGR1806–20, as well as complex multi-peaked time histories of 'classical' bursts might

instead be associated with interactions of the neutron star with separate fragments of the exploded parent small body. These possibilities will be checked in a further theoretical study. $\square$

Received 15 November 1989; accepted 15 January 1990.

1. Sagdeev, R. Z. & Zacharov, A. V. *Nature* **341**, 581–585 (1989).
2. Mitrofanov, I. & Barat, C. *Proc. Am. Inst. Phys. Conf. No. 170*, 293–301 (American Institute of Physics, New York, 1988).
3. Ruderman, M. & Cheng, A. *Astrophys. J.* **335**, 306–318 (1988).
4. Harwitt, M. & Salpeter, E. E. *Astrophys. J.* **186**, L37–L39 (1973).
5. Shklovskij, I. S. *Soviet Astr.* **51**, 665–667 (1974).
6. Colgate, S. A. & Petschec, A. G. *Astrophys. J.* **248**, 771–782 (1981).
7. Van Buren, D. *Astrophys. J.* **249**, 297–301 (1981).
8. Mitrofanov, I. G. Preprint No. 1597, Space Research Institute, Moscow (1989).
9. Marochnik, L. C., Mukhin, L. M. & Sagdeev, R. Z. *Science* **242**, 547–550 (1988).
10. Sagdeev, R. Z. *et al. Nature* **321**, 259–262 (1986).
11. Newman, M. G. & Cox, A. N. *Astrophys. J.* **242**, 319–325 (1980).
12. Tremaine, S. & Zytow, A. N. *Astrophys. J.* **301**, 155–163 (1986).
13. Weismann, P. R. *Astr. Astrophys.* **118**, 90–94 (1983).
14. Hurley, K. *Proc. Am. Inst. Phys. Conf. No. 141* (eds Liang, E. P. & Petrosian, V.), 1–74 (American Institute of Physics, New York, 1986).
15. Lamb, D. Q., Lamb, F. K. & Pines, D. *Nature* **246**, 52–54 (1973).
16. Mazets, E. P. & Golenetskii, S. V. Preprint No. 632, A. F. Ioffe Physical-Technical Institute (1979).
17. Golenetskii, S. V., Ilyinskii, V. N. & Mazets, E. P. *Nature* **307**, 41–42 (1984).
18. Atteia, J.-L. *et al. Astrophys. J.* **320**, L105–L110 (1987).
19. Boer, H., Hameury, J. M. & Lasota, J. P. *Nature* **337**, 716–718 (1989).
20. Allen, C. W. *Astrophysical Quantities* (Athlone, London, 1973).
21. Cheng, A., Ho, C. & Ruderman, M. *Astrophys. J.* **300**, 500–521 (1986).

Dynamics of Triton's atmosphere

Andrew P. Ingersoll

Division of Geological and Planetary Sciences,
California Institute of Technology, Pasadena, California 91125, USA

THE recent Voyager encounter established certain facts about Triton's atmosphere: the surface pressure is in the range 1.5–1.9 Pa (15–19 μ bar)¹; the surface temperature is 38 ± 3 K (ref. 2); molecular nitrogen is the dominant atmospheric constituent³; hazes and clouds are visible not only on the limb but also against the surface⁴; the wind in the southern hemisphere is to the north-east at low altitudes (as shown by streaks on the surface⁴) and to the west at high altitudes (as shown by geyser-like plume tails⁴). Triton rotates with a period of 5.877 days in a right-hand sense about the south pole, where the season now is late spring⁴. Here we argue that these features can be explained if Triton, like Mars⁵, has a global, well-structured atmosphere in equilibrium with surface frosts. The subliming frost cap produces a polar anticyclone at low altitudes, with northeastward winds of ~ 5 m s⁻¹ within the Ekman boundary layer. The temperature contrast between the cold frost-covered pole and the warm unfrosted equator produces westward winds at high altitudes.

Trafton⁶ and Trafton and Stern⁷ defined the difference between thick and thin atmospheres, both in vapour equilibrium with surface frosts. For thick atmospheres, the frost temperature is maintained at the global mean value through contact with the vapour, which covers the globe at nearly constant pressure. Sublimation requires a net outflow of gas from the summer hemisphere, but the winds are subsonic and do not require large pressure gradients (fractional variations of order unity) to sustain them. Ingersoll and coworkers^{8,9} have studied thin atmospheres, focusing on the atmosphere of Io. There, the winds reach their maximum possible speeds, several times the speed of sound, and the pressure varies by orders of magnitude from place to place. Warm frost near the sub-solar point and active volcanoes maintain local atmospheres, outside of which the pressure is essentially zero.

The Voyager science teams^{1–4} have analysed the Triton data in terms of a thick-atmosphere model, in which the albedo determines the global mean frost temperature which in turn determines the global mean pressure of N₂. Let F_a be $(1-A)F_0/\epsilon$, where A is the albedo of the frost, ϵ is its emissivity,

and F_0 is the incident solar flux (~ 1.52 W m⁻²) at the subsolar point. The global mean temperature in radiative equilibrium is given by $T = (F_a/4\sigma)^{1/4}$, where σ is the Stefan-Boltzmann constant and the factor of 1/4 is the ratio of cross-sectional area to surface area of a sphere. Both A and ϵ are uncertain^{4,10}, but for $\epsilon = 1.0$ and $A = 0.7$, the equilibrium temperature is 37.6 K. At this temperature the vapour pressure of N₂ is ~ 1.5 Pa (ref. 11), in good agreement with the Voyager measurements.

Because the frost temperature is constant, the net radiative heating (visible minus infrared) in the sunlit hemisphere goes entirely into sublimation. This fact implies a certain mass transport into the dark hemisphere. If the wind speed were independent of altitude, the outflow velocity, v_0 , near the terminator (the boundary between the day and night sides) would be of the order of $F_a g r / (4PL)$, where r is Triton's radius ($\sim 1,352$ km), g is the surface gravity (~ 0.79 m s⁻²), L is the latent heat of vaporization¹¹ ($\sim 2.5 \times 10^5$ J kg⁻¹), and P/g is the mass/area ratio of the atmosphere. The factors of 1/4 in the expressions for T and v_0 assume a frost-covered planet. The appropriate factors are smaller than 1/4 if the area of frost on the day side is less than that on the night side. Substituting $P = 1.5$ Pa, one obtains $v_0 = 0.32$ m s⁻¹, which is much less than the speed of sound (~ 125 m s⁻¹). The same calculation for Io^{8,9}, with $P = 10^{-2}$ Pa, yields $v_0 = 2,500$ m s⁻¹, which is much greater than the speed of sound and is therefore inconsistent with the assumption of global vapour equilibrium.

The above reasoning puts Triton in the same class as Mars, but the analogy is not perfect. The martian atmosphere contains ~ 3 –5 times as much mass as the seasonal polar caps. Averaged over the year, the caps neither gain nor lose net mass. Thus the annual average energy budget at the poles determines the temperature of the frost and its vapour pressure⁵. Triton's atmosphere contains much less mass than the seasonal frost deposits, which accumulate at a rate given by $\sigma T^4/L$. The depth of frost that accumulates during the 82-year autumn and winter seasons is ~ 1.2 m, assuming a density of 10^3 kg m⁻³. If the caps extend to a latitude of 45°, the seasonal deposits at each cap have 90 times the mass of the atmosphere. As on Mars, the atmosphere maintains all the frosts at the same temperature. But on Triton this temperature is determined by the instantaneous energy budget of all the frosts linked together, rather than the annual average energy budget of the permanent frost cap.

The estimate of the wind speed given earlier is based entirely on mass conservation and ignores the dynamics of the atmosphere. The calculated v_0 is, however, much less than Triton's tangential speed of rotation, $v_r = r\Omega = 16.7$ m s⁻¹, where Ω is the angular velocity of the solid planet. This fact means that Coriolis forces must come into play. The outward flow is deflected to the right (eastward) in the southern hemisphere, because Triton rotates in the opposite sense to the Earth. If the flow were geostrophic^{12,13}, so that Coriolis forces balance pressure gradients, the dominant flow would be a clockwise (anticyclonic) vortex centred at the south pole. The outflow would be confined to a thin Ekman layer near the surface, of thickness $d = \sqrt{\nu/\Omega}$, where ν is the kinematic viscosity of the gas^{12,13}. The wind speed in the Ekman layer would be of the same magnitude as that in the atmosphere above. To balance the sublimation outflow, both speeds would be larger than v_0 by a factor of order H/d , where H is the scale height, a measure of atmospheric thickness. Scaling from air at standard conditions and assuming that the molecular viscosity varies as $T^{1/2}\rho^{-1}$, where ρ is density, yields $\nu = 0.045$ m² s⁻¹, whence $d = 60$ m. The scale height H is given by RT/g , where R is the gas constant divided by the molecular weight of N₂, whence $H = 14.8$ km. This gives $H/d = 245$ and a wind speed, $v_0 H/d$, of ~ 78 m s⁻¹.

This second estimate of the wind speed is too high, for two reasons. One reason has to do with turbulence; the molecular viscosity underestimates the shear stress. The other reason has to do with the finite speed of rotation. The winds in a steady polar anticyclone cannot exceed v_r , because the circulation

would then be a cyclone blowing in the opposite direction, with inflow rather than outflow. The highest central pressure that a balanced polar vortex can sustain corresponds to that for a non-rotating atmosphere. Then the wind speed relative to the surface is $v_r \cos \phi$, where ϕ is the latitude. Such a wind would appear to blow to the east on Triton—opposite to the direction of rotation. This is the limiting wind speed of the polar anticyclone on Triton.

Turbulence probably reduces the wind speed below this limiting value. The Reynolds number $Re = vd/\nu$ based on the molecular viscosity is large, of order 2×10^4 when $v = v_r = 16.7 \text{ m s}^{-1}$, $d = 60 \text{ m}$ and $\nu = 0.045 \text{ m}^2 \text{ s}^{-1}$. A more appropriate Reynolds number is based on the eddy viscosity $\nu_e = u_*^2/(dv/dz)^{-1}$, where u_* is the friction velocity, defined so that ρu_*^2 is the surface stress¹⁴. With $dv/dz = v/d_e$, the Reynolds number becomes $Re = vd_e/\nu_e = v^2/u_*^2 = 1/C_D$, where C_D is the drag coefficient (~ 0.002 for typical turbulent boundary layers¹⁴). This means that $\nu_e = C_D v d_e = C_D v_0 H = 10 \text{ m}^2 \text{ s}^{-1}$; vd_e is related to the vertically integrated mass transport in the Ekman layer and is given by $v_0 H$, a known quantity for Triton. With $\nu_e = 10 \text{ m}^2 \text{ s}^{-1}$, the Ekman layer depth $d_e = \sqrt{\nu_e/\Omega}$ is $\sim 1 \text{ km}$, and the velocity in both the Ekman layer and in the atmosphere above is of order $v_0 H/d_e \approx 5 \text{ m s}^{-1}$. This estimate is highly uncertain, because the drag coefficient is uncertain by a factor of three or more. The important point is that the wind speed is equal to or less than v_r and much greater than v_0 .

The next step is to compare the predictions of the polar anticyclone model with observations of wind streaks and plume drifts⁴. At the surface, the wind speed is zero. In a laminar Ekman layer, the surface stress and the derivative of the wind with altitude are 45° to the left of the geostrophic wind (the wind above the boundary layer). For a polar anticyclone in Triton's southern hemisphere, this surface stress is to the north-east. The fact that the surface streaks are also to the north-east implies that they are controlled by processes in the boundary layer.

French and Gierasch¹⁵ applied the same reasoning to the polar vortex of Mars. As on Triton, the wind streaks spiral outwards from the subliming polar cap. In a turbulent Ekman layer the turning angle between the surface stress and the zonal wind depends on a parameter $B = C_D v/(\Omega \nu_e)^{1/2}$ (ref. 15). The above discussion of turbulence on Triton is equivalent to the assumption that $B = 1$. The actual value of B depends on the stability of the atmosphere, which is unknown for Triton. The best guess is that the turning angle is significantly different from zero but less than the 45° of a laminar Ekman layer.

The anticyclone is supported by an excess pressure at the pole of limiting magnitude βP , where $\beta = h/H$, and $h = r^2 \Omega^2/g$ is the difference between the equatorial and polar radii of Triton, because the atmosphere rotates slowly or not at all. The numbers given earlier yield $\beta = 0.024$, whence the pressure excess at the pole is a small fraction of P itself.

The temperature at the pole is controlled by vapour equilibrium. If the equator is free of frost, its temperature is controlled by radiative equilibrium, and the equator may be warmer than the pole. A similar situation occurs during the martian spring, when the polar frost depresses temperatures over the pole and allows large temperature gradients to develop at the cap edge¹⁶. The Voyager images suggest that the equator of Triton is free of frost⁴, implying that the temperatures are above the frost point. Moreover, the diurnally averaged solar heating at the equator is comparable to that for the planet as a whole. A slightly lower albedo for the equator then makes it warmer than the frost, for which the temperature is $(F_a/4\sigma)^{1/4}$ as explained earlier. These statements apply at the time of the Voyager encounter, when the sub-solar latitude was -45° .

The thermal-wind equation^{12,13} then relates the equatorial temperature excess ΔT to the increase of zonal wind with altitude. The equation indicates that the westward velocity at altitude z is greater than that at the surface by an amount of

order $(gz/2\Omega r)\Delta T/T$. For example, with $z = 8 \text{ km}$ and $\Delta T/T = 0.1$, this excess westward velocity is $\sim 19 \text{ m s}^{-1}$. Haberle *et al.*¹⁶ have used similar arguments to calculate winds in the atmosphere of Mars.

These arguments help to explain the behaviour of the geyser-like plume at -50° latitude⁴. It rises almost vertically to an altitude of 8 km , and then streams off horizontally to the west for a distance of 100 km . This is strong evidence that the polar atmosphere is colder than the atmosphere at the equator. According to this model, the winds blow to the north-east in the boundary layer, to the east above the boundary layer, and to the west at higher altitudes. The fact that the eastward component is not apparent in the images suggests that the vertical velocity in the plume is larger than the eastward horizontal velocity ($\sim 5\text{--}15 \text{ m s}^{-1}$) according to the present model.

These wind speeds are comfortably above the minimum required to transport particles over distances of 100 km or more in Triton's atmosphere. C. Sagan and C. Chyba (manuscript in preparation) estimate that particles of $5 \mu\text{m}$ diameter will fall 1 km as they travel 100 km in a wind of 4.5 m s^{-1} . The horizontal distance scales inversely with particle diameter and linearly with wind speed and fall distance.

The above discussion ignores the diurnally varying parts of the circulation (tides), which could arise from absorption of sunlight in the atmosphere or from the daily sublimation of surface frosts. Diurnal heating of the atmosphere produces a daily temperature fluctuation of order αT , where $\alpha = 1/(\Omega t_R)$, $t_R = 4PC_p T/(gF_0\tau)$ is a radiative time constant, τ is the visible optical thickness of the atmosphere and C_p is the heat capacity of the frost. The corresponding pressure fluctuation is αP , where $\alpha = 0.004$ for $\tau = 10^{-2}$. Both of these numbers are upper limits⁴, so the diurnal pressure fluctuation associated with atmospheric heating is likely to be small compared with βP , the pressure difference associated with the polar anticyclone when $\beta = 0.024$.

There is also a daily pressure oscillation associated with frost on the surface that goes in and out of the sunlight. If the heat capacity of the frost is negligible, the peak sublimation rate is $F_a \cos \zeta/L$, where ζ is the solar zenith angle. Multiplying by g gives the rate of change of P , and dividing by Ω gives the peak daily variation of P . Setting the average of $\cos \zeta$ equal to $1/4$ gives a pressure variation of order ηP , where $\eta = v_0/v_r = 0.019$. This diurnal pressure variation is a non-negligible fraction of βP . If, however, the sunlight is absorbed over a depth d_s , the daily temperature variation of the frost is of order $F_a/(4\rho_s d_s C_p \Omega)$, where ρ_s is the density of the frost. For $d_s = 1 \text{ m}$ and $C_p = 1,600 \text{ J kg}^{-1} \text{ K}^{-1}$, this temperature variation ΔT is $\sim 0.006 \text{ K}$. The temperature variation in the frost produces a pressure variation of magnitude δP , where $\delta = L\Delta T/RT^2 = 0.003$ when $\Delta T = 0.006 \text{ K}$. This pressure rise is much less than βP , indicating that 1 m of frost significantly dampens the diurnal pressure variation. In fact, spectroscopic studies^{17,18} suggest that sunlight penetrates deeper than 1 m , so the diurnal circulation is likely to be small compared to the zonally symmetric circulation.

The main conclusion is that the winds are between 5 and 15 m s^{-1} . Wind direction is to the north-east near the ground, then to the east, and finally to the west at high altitudes. The change from east to west follows from the thermal-wind equation, with the atmosphere over the illuminated pole being colder than the atmosphere over the equator.

Received 16 November 1989; accepted 15 February 1990.

1. Tyler, L. *et al. Science* **246**, 1466–1473 (1989).
2. Conrath, B. *et al. Science* **246**, 1454–1459 (1989).
3. Broadfoot, L. *et al. Science* **246**, 1459–1466 (1989).
4. Smith, B. A. *et al. Science* **246**, 1422–1449 (1989).
5. Leighton, R. B. & Murray, B. C. *Science* **153**, 136–144 (1966).
6. Trafton, L. *Icarus* **58**, 312–324 (1984).
7. Trafton, L. & Stern, S. A. *Astrophys. J.* **267**, 872–881 (1983).
8. Ingersoll, A. P., Summers, M. E. & Schlipt, S. G. *Icarus* **64**, 375–390 (1985).
9. Ingersoll, A. P. *Icarus* **81**, 298–313 (1989).

10. Lane, A. L. *et al. Science* **246**, 1450–1454 (1989).
11. Brown, G. N. Jr & Ziegler, W. T. *Adv. Cryogen. Engng* **25**, 662–670 (1980).
12. Holton, J. R. *An Introduction to Dynamic Meteorology* 2nd Edn (Academic, New York, 1979).
13. Houghton, J. T. *The Physics of Atmospheres* 2nd Edn (Cambridge University Press, 1986).
14. Priestley, C. H. B. *Turbulent Transfer in the Lower Atmosphere* (University of Chicago Press, 1959).
15. French, R. G. & Gierasch, P. J. *J. geophys. Res.* **84**, 4634–4642 (1979).
16. Haberle, R. M., Leovy, C. B. & Pollack, J. B. *Icarus* **39**, 151–183 (1979).
17. Cruikshank, D. P., Brown, R. H. & Clark, R. N. *Icarus* **58**, 293–305 (1984).
18. Cruikshank, D. P., Brown, R. H., Giver, L. P. & Tokunaga, A. T. *Science* **245**, 283–286 (1989).

ACKNOWLEDGEMENTS. I thank R. H. Brown, C. Chyba, C. Leovy, S. Leroy, R. M. Nelson, C. Sagan, J. Spencer, E. Stone, D. Strobel, R. Thompson, R. Yelle and R. Zurek for useful comments and discussions. The research was supported by the NASA Planetary Atmospheres Program and by Voyager Project funds.

Microscopy of chemical-potential variations on an atomic scale

C. C. Williams & H. K. Wickramasinghe*

IBM T. J. Watson Research Center, PO Box 218, Yorktown Heights, New York, New York 10598, USA

* To whom correspondence should be addressed

THE invention of the scanning tunnelling microscope¹ (STM) has stimulated the development of several new forms of probe microscopy^{2–10}. Here we demonstrate the use of a microscope that is capable of measuring chemical-potential variations on an atomic scale—the scanning chemical potential microscope (SCPM). The system is based on a recently developed tunnelling thermometer¹¹, which allows the spatial mapping, on an atomic scale, of thermoelectric potential variations resulting from absorption of light, by scanning a conducting tip within tunnelling range of a conducting (or semiconducting) sample. In the SCPM, we replace the optical pump with an electrical sample heater, to generate a temperature gradient between the sample and the tunnel-current-measuring device. We measure the spatial variations in the thermoelectric

voltage across the tip–sample system as the tip is scanned across the sample surface with no external bias. This signal can be shown to be equal to the product of the local gradient of chemical potential with respect to temperature and the temperature differential normal to the surface being imaged. The images obtained in this way show features that are not present in the conventional STM images.

The measurement of chemical-potential gradients can be understood in the following way. As the probe tip and sample come within tunnelling range (5 Å or less), their electronic states become strongly coupled and their chemical potentials tend to equalize by two-way electron tunnelling across the gap. If there is a temperature gradient between the tunnel-junction region and the device that measures the tunnel current, a thermoelectric voltage will be set up across the tip–sample junction because of differential variations in chemical potential with temperature of the tip and sample. If T is the local equilibrium temperature at the junction, and T_0 is the temperature of the leads at the measurement end, the measured thermoelectric potential would be

$$\int_{T_0}^T \left(\frac{\partial \mu_t}{\partial T} - \frac{\partial \mu_s}{\partial T} \right) dT$$

where μ_s and μ_t are the chemical potentials of the sample and tip respectively. For small temperature differentials $T - T_0 = \Delta T_1$, the measured signal will be $[\mu'_t(T)\Delta T_1 - \mu'_s(T)\Delta T_1]$, where the prime represents gradients with respect to temperature. The variations that we measure as we scan the tip across the sample will appear as a small modulation on top of this background thermoelectric signal. If the chemical-potential gradient of the tip (μ'_t) remains constant throughout the scans, the only signal that will contribute to variations over small spatial scales will be that part of the thermoelectric potential generated by temperature gradients across the surface layer being imaged. Thus the thermoelectric signal will vary as $(\partial \mu_s(x, y)/\partial T)\Delta T$, where ΔT is the temperature differential normal to the surface layer.

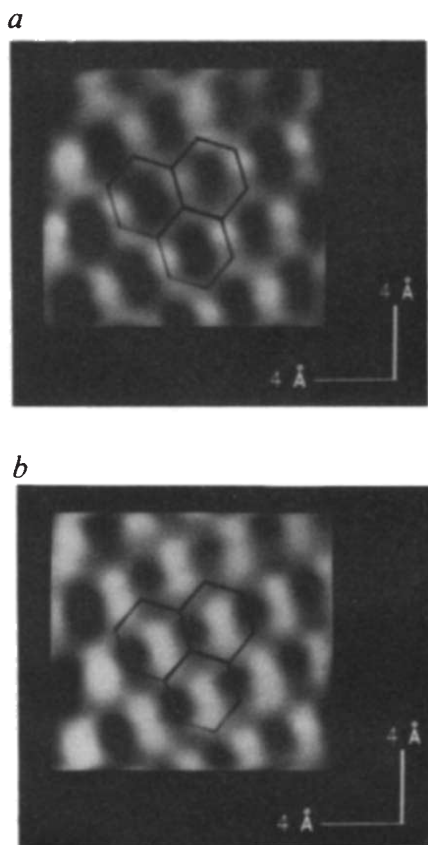


FIG. 1 a, STM image and b, chemical-potential-gradient image of MoS₂.

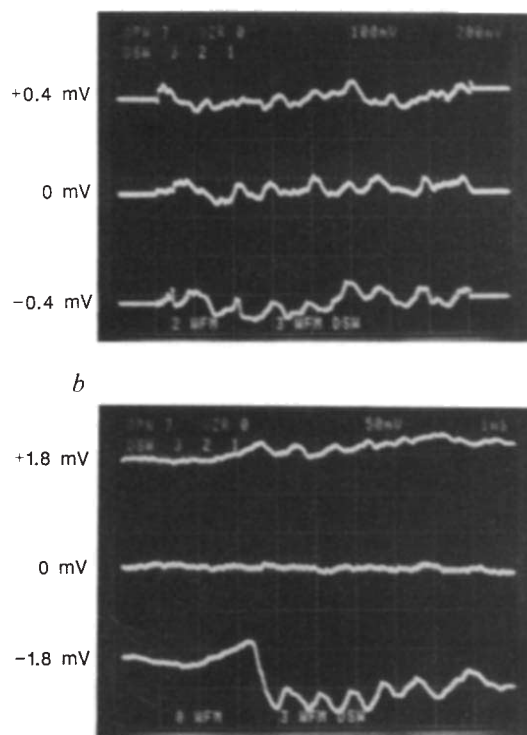


FIG. 2 Line scans in constant-height mode of the signal variations across a MoS₂ surface. The top trace is for a positive bias, the middle trace for zero bias and the bottom trace for negative bias. Scan field is 35 Å and peak-to-peak signal variations are ~0.5 mV. a is for a surface heated to 10 K above ambient whereas b was recorded at ambient temperature.

The recorded image is therefore proportional to variations in chemical-potential gradients across the sample.

In our initial experiments, we used a system similar to that described in ref. 11, but instead of alternating between the normal STM mode and the 'chemical-potential-gradient' mode at each pixel, we did so for every other line in the image. We manipulated the data by computer so as to obtain two separate images (chemical-potential gradient and STM) in registration with one another. The images were taken with the loop open and in the constant-height mode. The frame rate of the images in both STM and chemical-potential-gradient modes was one frame per second, and the video bandwidth was 30 kHz.

In STM mode, voltages were applied to the tip through a 100-M Ω bias resistor and the sample was connected to ground. The voltage at the sample was monitored using an amplifier with an input impedance of 100 M Ω and a gain of 100, which recorded the variations in gap impedance in the constant-height mode. In chemical-potential-gradient mode, the bias voltage was reduced to zero, thereby measuring only the thermoelectric voltage across the gap. The gap impedance was always much less than the STM bias resistance so that the chemical-potential gradient signal was not loaded by the bias resistance. Furthermore, we applied a nulling bias voltage to the amplifier to offset any d.c. thermoelectric voltages that may have been present in the system, including the d.c. thermoelectric contribution from the tip-sample junction.

The sample was attached to a small heater stage which was capable of heating it up to 50 K above ambient. We ran our experiments with the sample heated to typically 25 K above ambient.

Our initial experiments were performed with the layered semiconductor MoS₂ and an etched platinum tip. In Fig. 1a we show the regular STM image of the surface taken with a bias voltage of 10 mV and a current of 1 nA. This image shows a hexagonal

atomic arrangement similar to what has been reported previously^{12,13}. Figure 1b shows the corresponding chemical-potential-gradient image, which was recorded simultaneously following the procedure above. This image shows strong signals where the tunnelling signal shows weak ones, and vice versa. Spatial variations in chemical-potential gradient across the sample are seen with atomic resolution. To check that we were indeed recording chemical-potential gradients, we performed further single-line scans across the surface at 30 Hz, adjusting the bias voltage between successive line scans. The results are shown in Fig. 2a. The top trace was taken with an average bias of +0.4 mV, and the bottom trace with an average bias of -0.4 mV. The middle trace shows a scan with an average bias voltage of zero; atomic corrugations are still clearly visible. We then turned off the heater and repeated the measurement (Fig. 2b). The top and bottom traces still show the corrugation but the middle trace shows no structure. This is evidence that the SCPM is indeed monitoring changes in surface chemical-potential gradients.

In Fig. 3 we show images obtained of a cleaved graphite surface. Again, the hexagonal structure is clearly visible in the STM image (Fig. 3a), taken with a bias voltage of 10 mV and a current of 1 nA. Figure 3b shows the chemical-potential-gradient image, which again exhibits a large signal where the STM signal is small. The signal variations here were ~ 20 μ V. Finally, we have imaged a graphite surface coated with a film of the liquid crystal PYP909. Figure 4a shows the STM image and Fig. 4b shows the corresponding chemical-potential-gradient image. Again, there are features in the SCPM image that are not present in the STM image, and vice versa.

By recording a series of images at different sample temperatures, it should be possible (with *a priori* knowledge of the variation of μ' with T) to distinguish between the atomic constituents of a surface. We believe that with further refinement of our system—such as independent feedback control for the

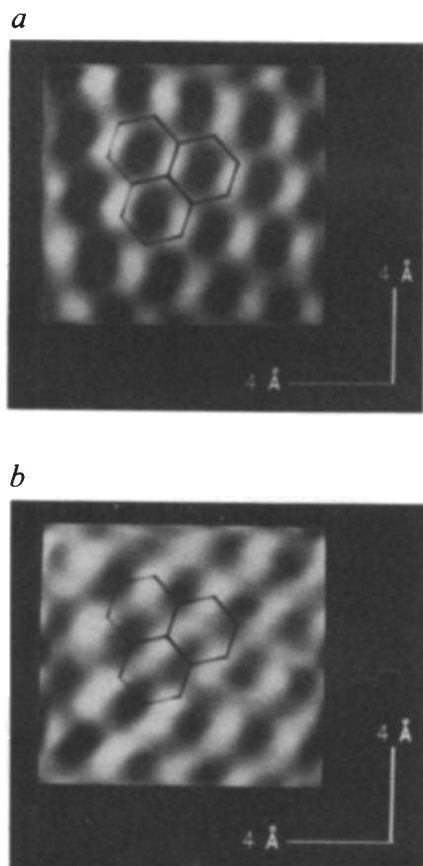


FIG. 3 a, STM image and b, chemical-potential-gradient image of graphite.

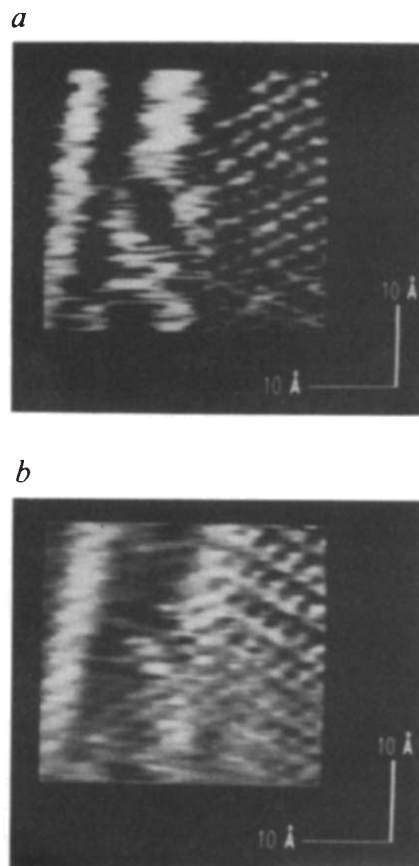


FIG. 4 a, STM image and b, chemical-potential-gradient image of liquid crystal PYP909 on graphite.

gap and more sensitive voltage measurements—the SCPM will have many important applications. For example, it might be used to study *in situ* chemical reactions at surfaces on an atomic scale. The ability to measure variations in chemical potential also allows the possibility of selectively identifying subunits of biological macromolecules either through a direct measurement of their chemical-potential gradients or by decorating them with different metals. This suggests a potentially simple method for sequencing DNA.

While no explicit calculations of chemical-potential gradients and tunnel-current probabilities have been made for these materials, it is not surprising that the maxima for the chemical-potential-gradient signal does not overlap strongly with the tunnel-current maxima. One major reason is the fact that the tunnelling current depends exponentially on the tip-sample gap, whereas the chemical-potential-gradient signal does not depend on the gap unless the gap impedance becomes comparable to the preamplifier resistance. A second reason for the differences is that the chemical-potential gradient should not depend uniquely on the local density of states at the surface; rather, it should depend on the local gradient with temperature of the average electron potential. □

Received 9 August 1989; accepted 2 February 1990.

1. Binnig, G., Rohrer, H., Gerber, Ch. & Weibel, E. *Phys. Rev. Lett.* **49**, 57–61 (1982).
2. Durig, U., Pohl, D. W. & Rohrer, F. *J. appl. Phys.* **59**, 3318–3327 (1986).
3. Betzig, E., Lewis, A., Harootunian, A., Isaacson, M. & Kratschmer, E. *Biophys. J.* **49**, 269–279 (1986).
4. Williams, C. C. & Wickramasinghe, H. K. *Appl. Phys. Lett.* **49**, 1587–1589 (1986).
5. Binnig, G., Quate, C. F. & Gerber, C. *Phys. Rev. Lett.* **56**, 930–933 (1986).
6. Martin, Y. & Wickramasinghe, H. K. *Appl. Phys. Lett.* **50**, 1455–1457 (1987).
7. Martin, Y., Abraham, D. W. & Wickramasinghe, H. K. *Appl. Phys. Lett.* **52**, 1103–1105 (1988).
8. Mate, C. M., McClelland, G. M., Erlandson, R. & Chiang, S. *Phys. Rev. Lett.* **59**, 1942–1945 (1987).
9. Manassen, Y., Hamers, R. J., Demuth, J. E. & Castellano, A. E. Jr. *Phys. Rev. Lett.* **62**, 2531–2534 (1989).
10. Hansma, P. K., Drake, B., Marti, O., Gould, S. A. C. & Prater, C. B. *Science* **243**, 641–643 (1989).
11. Weaver, J. M. R., Walpita, L. M. & Wickramasinghe, H. K. *Nature* **342**, 783–785 (1989).
12. Stupian, G. W. & Leung, M. S. *Appl. Phys. Lett.* **51**, 1560–1562 (1987).
13. Sarid, D. & Henson, D. *Appl. Phys. Lett.* **52**, 2252–2254 (1988).

ACKNOWLEDGEMENTS. We thank D. Sarid for providing the MoS₂ sample, J. M. R. Weaver for helpful discussions, and D. W. Abraham and J. J. Matrangola for help with the image processing.

Selective oxidation of methane to synthesis gas using transition metal catalysts

A. T. Ashcroft*, A. K. Cheetham*, J. S. Foord†, M. L. H. Green†, C. P. Grey*, A. J. Murrell† & P. D. F. Vernon†

* Chemical Crystallography Laboratory, University of Oxford, 9 Parks Road, Oxford OX1 3PD, UK

† Inorganic Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QR, UK

THE diminishing reserves of petroleum oil have focused attention on the possibility of making more efficient use of natural gas, reserves of which are at present considerably under-utilized. Methane is commonly used as a fuel, but it is also the starting material for the production, by steam reforming, of synthesis gas (carbon monoxide and hydrogen), which acts as a feedstock for the synthesis of ammonia and methanol, and can be converted to higher hydrocarbons, alcohols and aldehydes by Fischer-Tropsch catalysis¹. The partial oxidation of methane to synthesis gas is also an established industrial process² but operates at very high temperatures (>1,200 °C). Here we report that this reaction can be carried out at temperatures of only ~775 °C by using lanthanide ruthenium oxide catalysts.

Attempts to convert methane directly into more valuable chemicals have focused on oxidative coupling reactions to yield ethylene and ethane^{3,4} and direct oxygenation to methanol and

TABLE 1 Results of catalytic reactions at 777 °C in the presence of nitrogen diluent

Catalyst	% Methane converted	% CH ₄ converted to CO	H ₂
RuO ₂	60	81	76
Pr ₂ Ru ₂ O ₇	94	98	99
Sm ₂ Ru ₂ O ₇	89	94	98
Eu ₂ Ru ₂ O ₇	92	97	99
Gd ₂ Ru ₂ O ₇	93	97	99
Tb ₂ Ru ₂ O ₇	89	94	96
Dy ₂ Ru ₂ O ₇	93	96	98
Tm ₂ Ru ₂ O ₇	90	96	98
Yb ₂ Ru ₂ O ₇	92	97	99
Lu ₂ Ru ₂ O ₇	94	97	98
1% Ru/Al ₂ O ₃ *	94	97	99
0.1% Ru/Al ₂ O ₃ *	61	78	83
No catalyst†	4	40	0

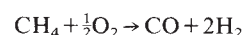
Reactant gas flow rates: N₂, 12 ml min⁻¹; CH₄, 6 ml min⁻¹; O₂, 3 ml min⁻¹. Total GHSV (Gas Hourly Space Velocity), 4 × 10⁴ h⁻¹; pressure, 1 atm. All experiments were carried out using 50 mg of solid, powdered catalyst, lightly packed between <20 mg of silica wool (Multilab) in a straight silica reaction tube of inner diameter ~4 mm. The reaction tube (300 mm) was placed in the vertical tube furnace of a Labcon microreactor and connected to a supply of the gas reaction mixture. The reactant gases, methane (Union Carbide, Gas and Equipment Ltd), oxygen (Air Products) and nitrogen (Air Products) were dried over molecular sieves and passed over the catalyst at a rate of 20–50 ml min⁻¹ (GHSV of 4–7 × 10⁴ h⁻¹). The temperature of the reaction tube was raised from ambient to the required temperature (typically 777 °C) over a period of 2 h. The reaction products were monitored using an on-line Hewlett-Packard 5890A gas-chromatography apparatus. Separation of all gases was obtained using helium carrier gas through Porapak Q and 5-Å molecular-sieve packed columns, and were detected using a thermal conductivity detector, calibrated on site. In all cases, O₂ conversion was >99.5%, and C, H, O, N mass balances were better than 96%.

* Prepared from RuCl₃ (anhydrous; Aldrich).

† The other principal products were CO₂ (24%), C₂H₄ (13%) and C₂H₆ (23%).

formaldehyde^{5,6}. Unfortunately, under conditions where reactions of methane are fast enough to be of interest (typically >700 °C), the formation of CO₂ is so favourable (ΔG < -800 kJ mol⁻¹) that partial oxidation to more useful products is difficult to achieve on an economical scale.

We have recently shown that a number of rare-earth tin oxides of the pyrochlore type, Ln₂Sn₂O₇ (where Ln is a lanthanide), are good catalysts for the oxidative coupling of methane to ethylene⁷. We now demonstrate that the ruthenium analogues, Ln₂Ru₂O₇, are extraordinarily active and selective catalysts for the partial oxidation of methane to synthesis gas according to the reaction:



Synthesis gas can be made by a number of methods, most of which involve the steam reforming of hydrocarbons or coal⁸. Methane, for example, can be converted over a nickel/alumina catalyst⁹ at 700–800 °C according to CH₄ + H₂O → CO + 3H₂. This reaction is an important source of carbon monoxide and hydrogen, but it is highly endothermic, and leads in addition to the formation of carbon dioxide by the shift reaction CO + H₂O = CO₂ + H₂. The partial oxidation reaction, by contrast, is mildly exothermic, more selective, and yields an H₂/CO ratio that is lower than that obtained by steam reforming. This lower ratio may be highly desirable for certain applications of synthesis gas. Indeed, secondary reformers using CO₂ or O₂ oxidants are frequently required to reduce the hydrogen content of synthesis gas made by steam reforming¹⁰.

Polycrystalline samples of Ln₂Ru₂O₇ were prepared at 1,100 °C by conventional solid-state reactions between Ln₂O₃ and RuO₂ in a sealed silica tube¹¹. The samples were characterized by X-ray powder diffraction (XRD), analytical electron microscopy (AEM), X-ray photoelectron spectroscopy (XPS)

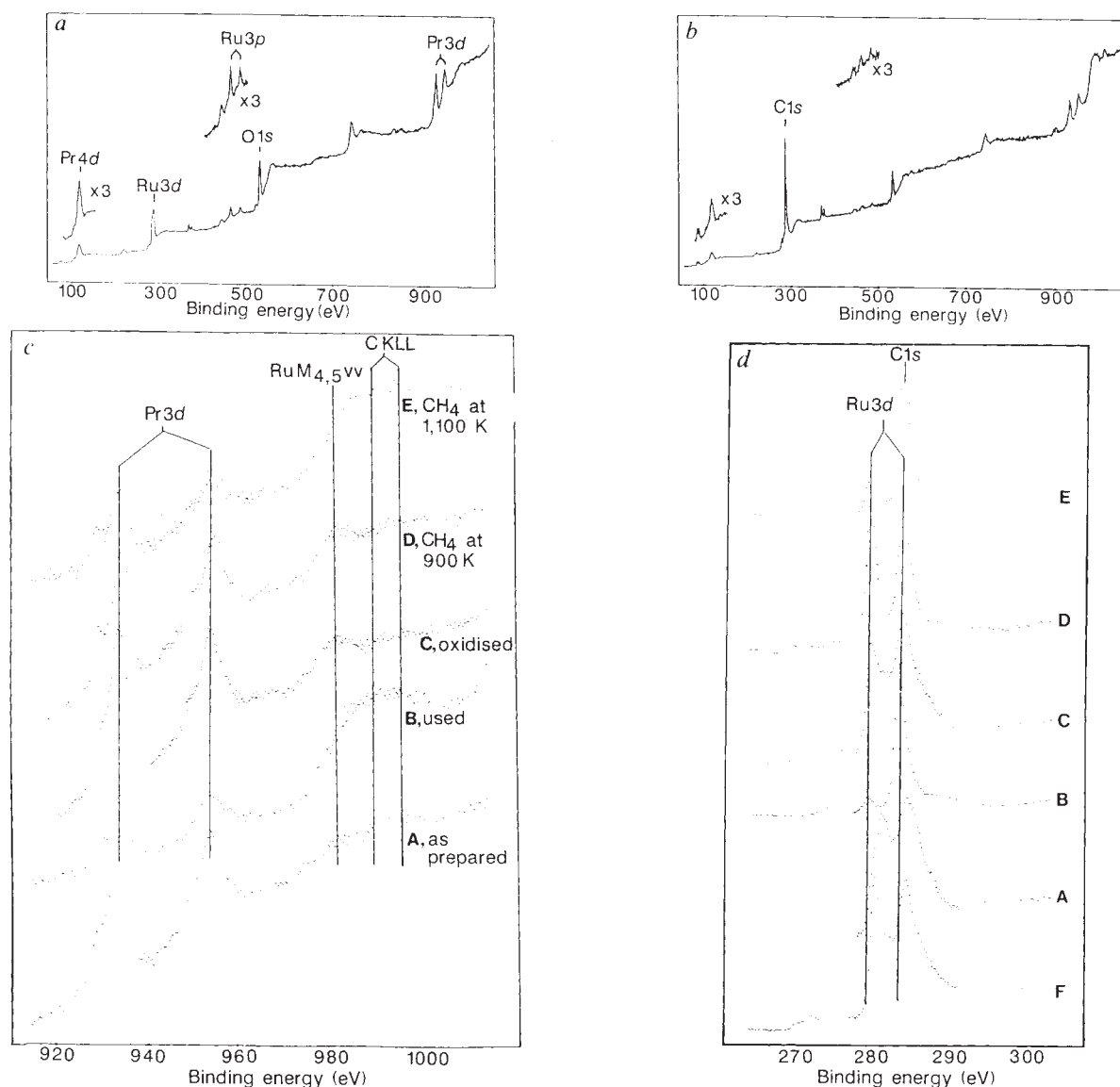


FIG. 1 Typical Mg K α XPS data, recorded with fixed analyser transmission (FAT). The six samples studied were: A, Pr₂Ru₂O₇ formed by heating PrO_x and RuO₂ at 900 °C for 100 h in air; B, the used Pr₂Ru₂O₇ catalyst; C, the material formed by heating B *in situ* in pure O₂ (100 mbar) at 620 °C for 10 min; D, the material formed by heating C *in situ* in methane (100 mbar) at 620 °C for 10 min; E, the material formed by heating D *in situ* in methane

(100 mbar) at 820 °C for 10 min; F, RuO₂ (Aldrich Chemical Co. Ltd) as used in the catalyst syntheses. *a*, A low-resolution spectrum of an initial sample of Pr₂Ru₂O₇. *b*, A low-resolution spectrum of a sample of Pr₂Ru₂O₇ after use in a catalytic reaction. *c*, A high-resolution spectrum of samples A–E of Pr₂Ru₂O₇, showing the Pr 3d, Ru M_{4,5} and C KLL peaks. *d*, A high-resolution spectrum of samples A–F, showing the Ru 3d and C 1s peaks.

and high-resolution transmission electron microscopy (HRTEM). The catalytic activity was examined under a variety of conditions, as detailed in Table 1 (for a feedstock containing CH₄:O₂:N₂ in molar ratios 2:1:4) and Table 2 (CH₄:O₂ molar ratio 2:1). The reaction with the Pr₂Ru₂O₇ catalyst was also examined at a total pressure of 20 atmospheres (Table 2). In reactions with the nitrogen diluent, methane conversions in excess of 90%, with selectivities in the range 94–99%, were found for all of the rare-earth pyrochlore phases after an induction period lasting approximately 30 min. Pure ruthenium dioxide was found to be a less effective catalyst. In the absence of the nitrogen diluent, conversions were rather lower (80–90%) with selectivities of about 95%; this is presumably due to the greater partial pressure of methane and oxygen. In longer experiments lasting up to seven days, no loss of activity or selectivity was observed.

We observe that the selectivity for synthesis gas increases with temperature. For example, the CO:CO₂ ratio increases from ~1:20 at 377 °C to >10:1 at 777 °C. Similarly, the corresponding increase in H₂:H₂O ratio is typically from ~1:4 to >20:1. Our

thermochemical calculations, based upon values for the heats and entropies of formation of CO and other products at 25 °C, indicate that the formation of synthesis gas becomes more favourable than CO₂ formation for $T > 693$ °C, for a CH₄:O₂ ratio of 2:1. At this stoichiometry, there is insufficient oxygen to convert all of the methane to CO₂, so that the formation of synthesis gas, which is entropically more favourable, is preferred at higher temperatures.

Examination of the used rare-earth catalysts by XRD, AEM, XPS and HRTEM revealed that partial reduction of the pyrochlores had taken place, leading to the formation of ruthenium metal. This was accompanied by a concomitant increase in the Ln:Ru ratio in the bulk mixed oxide, which now appeared to be a defective fluorite of the type (Ln, Ru)O_{2-x}.

Typical XPS data for five samples of Pr₂Ru₂O₇ (A–E) and for RuO₂ (F) are shown in Fig. 1. The experiments showed the following properties of the surfaces of samples A–E. (1) Over a sampling depth of ~20 Å the Pr:Ru ratio is approximately unity. However, for samples B, C, D and E, which have been reacted, the Pr:Ru ratio over the first 7 Å depth show enrichment

TABLE 2 Results of catalytic reactions at 777 °C in the absence of nitrogen diluent

Catalyst	% Methane converted	% CH ₄ converted to	
		CO	H ₂
Pr ₂ Ru ₂ O ₇	89	94	98
Sm ₂ Ru ₂ O ₇	74	85	90
Eu ₂ Ru ₂ O ₇	87	95	97
Gd ₂ Ru ₂ O ₇	87	94	97
Tb ₂ Ru ₂ O ₇	79	89	94
Dy ₂ Ru ₂ O ₇	86	94	97
Tm ₂ Ru ₂ O ₇	79	91	94
Yb ₂ Ru ₂ O ₇	82	91	95
Lu ₂ Ru ₂ O ₇	87	95	97
1% Ru/Al ₂ O ₃ *	89	95	98
Pr ₂ Ru ₂ O ₇ †	23	81	89
No catalyst‡	17	61	13

Reactant gas flow rates: CH₄, 14 ml min⁻¹; O₂, 7 ml min⁻¹. Total GHSV, 4 × 10⁴ h⁻¹; pressure, 1 atm.

* Prepared from RuCl₃ (anhydrous; Aldrich).

† Reaction carried out at 20 atm total pressure, with an initial CH₄/O₂ ratio of 6, to avoid the risk of explosion. Reduced conversion reflects excess of methane over stoichiometry; the O₂ conversion was still complete.

‡ The other principal products are CO₂ (7%), C₂H₄ (20%) and C₂H₆ (10%).

in ruthenium. Therefore, ruthenium enrichment of the surface occurs during the initiation of the catalyst. (2) Figure 1d shows that the peaks assigned to ruthenium in the surface of the samples that had been reacted with methane are all shifted to lower binding energies than for sample A. This suggests that the ruthenium is in a lower oxidation state in the activated catalyst. The Ru 3d peak occurs as a well-resolved doublet, which is characteristic of ruthenium metal rather than RuO₂ (ref. 12), suggesting that ruthenium on the surface has been reduced to the metal during reaction with methane. The Ru peak of C, which was reacted with oxygen, is analogous in appearance to that of RuO₂ (ref. 12). This surface layer can again be reduced to ruthenium metal by heating under methane, as in D. (3) The samples B and E, which have been exposed to methane at temperatures >720 °C (the temperature at which the reaction becomes more selective), showed substantial surface coverage of carbon (Fig. 1b, c, d). The binding energies for this carbon lie below those for carbon bonded to oxygen and above those normally associated with metal carbide species¹³, but are typical of those in a thick layer of graphitic carbon.

These results suggest that the active catalyst may be supported Ru metal, a view that is reinforced by the performance of a Ru/Al₂O₃ catalyst (Tables 1, 2). The mechanism remains unclear, but the carbonaceous surface may be involved, as it is in Fischer-Tropsch synthesis¹⁴, in the metal-catalysed hydrogenation of carbon monoxide to methane¹⁵, and in steam reforming¹⁶. The reaction pathway may involve an initial conversion of some methane to CO₂ and steam, followed by a sequence of steam reforming and reverse water-gas shift reactions. This would be consistent with some preliminary observations indicating that supported Ni, Pd and Rh are also excellent catalysts for the partial oxidation of methane. □

12. Kim, K. S. & Winograd, N. *J. Catal.* **35**, 66–72 (1972).

13. Ramqvist, L., Hamrin, K., Johansson, G., Fahlmann, A. & Nordling, C. *J. Phys. Chem. Solids* **30**, 1835–1847 (1969).

14. Rabo, J. A., Risch, A. P. & Poutsma, H. L. *J. Catal.* **53**, 295–305 (1978).

15. Winslow, P. & Bell, A. T. *J. Catal.* **91**, 142–154 (1985).

16. Jackson, S. D., Thomson, S. J. & Webb, G. *J. Catal.* **70**, 249–263 (1981).

ACKNOWLEDGEMENTS. We thank the Gas Research Institute for financial support, P.D.F.V. and C.P.G. thank the BP Research Centre, Sunbury-on-Thames, and the SERC respectively, for studentships. We also thank G. Harris for the HRTEM measurements, P. A. Cox for use of his ESCALAB XPS spectrometer, N. Parkyn for some early experimental confirmation of this work and for useful discussions, and R. Burch for advice on the reaction pathway.

Observations of denitrification and dehydration in the winter polar stratospheres

D. W. Fahey*, K. K. Kelly*, S. R. Kawa*, A. F. Tuck*, M. Loewenstein†, K. R. Chan† & L. E. Heidt‡

* Aeronomy Laboratory, National Oceanic and Atmospheric Administration, Boulder, Colorado 80303, USA

† Ames Research Center, National Aeronautics and Space Administration, Moffett Field, California 94035, USA

‡ National Center for Atmospheric Research, Boulder, Colorado 80303, USA

WATER vapour and reactive nitrogen species are removed from the polar stratosphere in both hemispheres in the winter and early spring, probably by the sedimentation of aerosol particles containing water and nitric acid, which co-condense at low stratospheric temperatures. In the Antarctic in 1987, intense dehydration invariably accompanied intense denitrification. However, from the marked difference between water vapour concentrations in the two hemispheres for similar concentrations of reactive nitrogen, we deduce that in the Antarctic some dehydration may have occurred without denitrification. In the Arctic in 1989, despite higher temperatures than in the Antarctic, intense denitrification occurred but without intense dehydration. These results provide important constraints for the uncertain microphysics controlling the growth and sedimentation of aerosol particles affecting the removal. We argue that the Arctic denitrification can be explained by the selective growth and sedimentation of aerosol particles rich in nitric acid. Because reactive nitrogen species moderate the destruction of ozone by chlorine-catalysed reactions, by sequestering chlorine in reservoir species such as ClONO₂, the possibility of the removal of reactive nitrogen without dehydration should be allowed for in attempts to model ozone depletion in the Arctic. Indeed, denitrification along with elevated concentrations of reactive chlorine¹ observed in 1989 indicate that the Arctic was chemically primed for ozone destruction without an extended period of temperatures below the frost point, as is characteristic of the Antarctic.

On-board measurements of reactive nitrogen, NO_y, and water vapour, H₂O, were obtained using the NASA ER-2 aircraft on flights poleward of 59° latitude at an altitude between 15 and 20 km (ref. 2). The missions occurred in austral late winter/early spring 1987 and in boreal mid and late winter 1989. Principal NO_y species are given in the sum

$$\text{NO}_y = \text{NO} + \text{NO}_2 + \text{NO}_3 + \text{HNO}_3 + 2(\text{N}_2\text{O}_5) + \text{HO}_2\text{NO}_2 + \text{ClONO}_2 \quad (1)$$

where the higher oxides are catalytically converted to NO followed by the detection of NO with an NO/O₃ chemiluminescence detector³. Measurements of H₂O are made through photodissociation by Lyman-α radiation to produce excited OH molecules and the subsequent detection of OH fluorescence⁴.

Figure 1 shows the qualitative signature of denitrification and dehydration, with NO_y and H₂O data from the Antarctic and Arctic, plotted as a function of latitude within the lower

Received 15 November 1989; accepted 8 February 1990.

- Henrici-Olivé, G. & Olivé, S. *Angew. Chem. Int. Ed. Eng.* **15**, 136–141 (1976).
- Kirk, R. E. & Othmer, D. F. (eds) *Encyclopedia of Chemical Technology* 3rd Edn, Vol. 12, 952 (Wiley Interscience, New York, 1980).
- Keller, G. E. & Bhasin, M. M. *J. Catal.* **73**, 9–19 (1982).
- Hutchings, G. J., Scurrell, M. S. & Woodhouse, J. R. *Chem. Soc. Rev.* **18**, 251–283 (1989).
- Gesser, H. D., Hunter, N. R. & Prakash, C. B. *Chem. Rev.* **85**, 235–244 (1985).
- Spencer, N. D. & Pereira, C. J. *J. Catal.* **116**, 399–406 (1989).
- Ashcroft, A. T., Cheetham, A. K., Green, M. L. H., Grey, C. P. & Vernon, P. D. F. *J. Chem. Soc. Chem. Commun.* **21**, 1667–1669 (1989).
- Keim, W. (ed.) *Catalysis in C₁ Chemistry* (Reidel, Dordrecht, 1983).
- Rostrup-Nielsen, J. R. in *Catalysis, Science & Technology* Vol. 5 (eds Anderson, J. R. & Boudart, M.) (Springer, Berlin, 1984).
- Topp-Jørgensen, J. in *Methane Conversion* (eds Bibby, D. M., Chang, C. D., Howe, R. F. & Yurchak, S.) 293 (Elsevier, Amsterdam, 1988).
- Bertaut, E. F., Forrat, F. & Montmory, M. C. *Compt. Rend. (Paris)* **249**, 829–831 (1959).

stratosphere. N_2O , an inert tracer in the lower stratosphere, serves as a reference for changes in air mass. The latitude gradients in N_2O are attributed to diabatic descent that increases poleward, bringing down air from higher in the stratosphere⁴⁻⁶. The boundary of the polar vortex is marked in each data set as defined chemically by elevated concentrations of the ClO radical (Antarctic), and dynamically by the maximum in the horizontal wind (Arctic)^{7,8}. Poleward of the boundary in the Antarctic, large decreases in both NO_y and H_2O can be seen, reversing the gradient in mixing ratio with latitude and N_2O , and signifying removal. In the Arctic, decreases occur further inside the vortex boundary and, in the case of H_2O , are smaller.

To quantify the removal of NO_y , the relation between N_2O and NO_y equatorward of the intensely denitrified region is used. N_2O and NO_y show a striking negative correlation in both hemispheres over a wide range of values of both species^{6,9}. The negative correlation is a result of the photochemical destruction of N_2O producing NO_y at low latitudes in the mid-stratosphere and subsequent transport of air poleward and downward. The average relation between N_2O and NO_y (see Fig. 1 legend) is adopted as a reference state which defines the quantity NO_y^* as a measure of unperturbed NO_y . Inside the vortex, the difference between observed NO_y and calculated NO_y^* values (see Fig. 1) provides a quantitative measure of denitrification⁶. A somewhat different approach must be used to quantify the removal of H_2O . The correlation between H_2O and N_2O is also negative in the unperturbed stratosphere, but the relation is not as consistent as that of NO_y with N_2O (ref. 4). A general correlation results because H_2O increases above the hygropause due to the oxidation of methane (CH_4), whereas N_2O and CH_4 decrease because of photochemical destruction. As N_2O decreases poleward at constant potential temperature, the amount of H_2O at the vortex boundary sets a lower limit to compare with values further poleward in the absence of dehydration. Differences of ≥ 0.5 parts per million by volume (p.p.m.v.) between observed H_2O values and this lower limit are taken to represent intense dehydration. Less intense dehydration can be detected by relative changes in residual water vapour, $R_{\text{H}_2\text{O}}$, inside the vortex. $R_{\text{H}_2\text{O}}$ is defined as the observed H_2O minus the contribution from CH_4 oxidation (see Fig. 3 legend)¹⁰. $R_{\text{H}_2\text{O}}$ represents the average water vapour content in the sampled air mass when it entered the stratosphere and is therefore assumed to be constant for individual flights near to and poleward of the vortex boundary⁴.

Using NO_y^* and $R_{\text{H}_2\text{O}}$, the quantitative correlation between the removal of NO_y and H_2O is shown in Fig. 2 for the flight portions shown in Fig. 1. The variations in the coupling of denitrification and dehydration are clearly seen. In the Arctic, at least 8 parts per 10^9 by volume (p.p.b.v.) of NO_y are removed

from air masses with an apparent average change in H_2O of <0.1 p.p.m.v. The largest reductions in NO_y , however, involve the removal of ~ 0.5 p.p.m.v. of H_2O in this case and in a later flight. In contrast, the Antarctic data show a large simultaneous removal of both species which suggests a strong coupling. Removal of up to 4 p.p.b.v. NO_y , however, shows little H_2O removal, as in the Arctic data. The intense denitrification observed in both data sets corresponds to a fractional removal (estimated using NO_y^*) as large as 90% of the available NO_y . The location inside the Antarctic vortex where removal occurs is probably downwind of the flight track (70°W)¹¹. Thus, values representing less than the maximum removal of both species in the Antarctic may represent mixing of air inside the vortex, where removal is more complete, with unperturbed air near the boundary⁶.

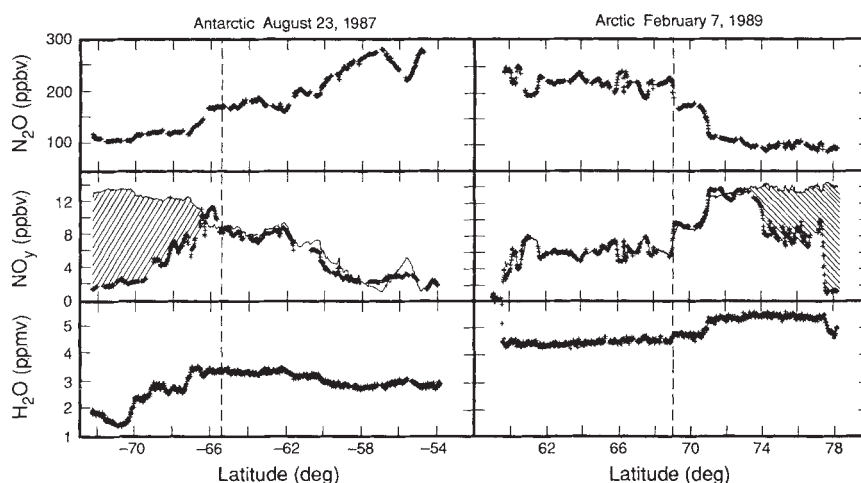
Figure 3 shows the scatter plots of $(\text{NO}_y - \text{NO}_y^*)$ and H_2O averaged over several flights where, by using as a reference an upper limit of measured N_2O , only data from inside the vortex are included in the average values. The average values show that the coupling found in Fig. 2 is systematic throughout the two data sets. The ordinate, H_2O , is chosen because of the difficulty in choosing representative $R_{\text{H}_2\text{O}}$ values from the Antarctic data⁴. With the H_2O scale, intense dehydration is still not found in the Arctic at any level of denitrification. The potential temperature range of 400–480 K in Fig. 3 corresponds to a geometric altitude of ~ 16 –20 km (ref. 12). When combined with the horizontal scales in Fig. 1 and the rapid circulation in the vortex, denitrification and, in the Antarctic, dehydration are shown to affect a large fraction of the polar vortex^{12,28}.

The coupling of the removal processes of NO_y and H_2O is through the formation and sedimentation of aerosol particles at low stratospheric temperatures¹³⁻¹⁵. The principal aerosol types are nitric acid trihydrate (NAT, type I, $\leq 1 \mu\text{m}$ diameter), formed by the co-condensation of H_2O and HNO_3 , and water ice (type II, $\geq 1 \mu\text{m}$ diameter), formed by the condensation of H_2O . An intermediate type II particle with HNO_3 as a dilute impurity may also play a role in the removal. Because HNO_3 is expected to be the largest component of the NO_y reservoir, removal of HNO_3 , by any process, results in intense denitrification in an air parcel. Data from both aircraft missions and balloon measurements have confirmed that HNO_3 is incorporated in particles at temperatures several degrees above the frost point, and that the conditions of formation match prediction based on theoretical and experimental studies (see, for example, refs 3, 16–20). As NAT condensate consists of HNO_3 and H_2O in a 1:3 mole ratio, the removal of stratospheric HNO_3 on sedimenting particles always includes some H_2O , but the amount need only be a small fraction of the available H_2O . The dehydration limit of 0.1 p.p.m.v. for denitrification of at least

FIG. 1 Data from a portion of a flight in the Antarctic and Arctic missions with individual points representing 10-s averages of the respective species. The data on 23 August are from the poleward leg, with potential temperatures increasing with latitude from 420 K to 450 K (~ 16 –19 km). The 7 February data are from the equatorward leg with potential temperatures between 450 K and 470 K (~ 17 –19 km). The average relation between NO_y and N_2O equatorward of the vortex boundary is given by

$$\text{NO}_y = -0.07(\text{N}_2\text{O}) + 21 \text{ p.p.b.v.} = \text{NO}_y^* \quad (2)$$

for N_2O values between 80 p.p.b.v. and 300 p.p.b.v. (ref. 8). The relation defines NO_y^* (solid line) as a measure of unperturbed NO_y poleward of the vortex boundary. The shaded areas, highlighting the difference between measured gas-phase NO_y and calculated NO_y^* , represent denitrification in the sampled air parcels. Note the relative magnitude of the mixing ratios of NO_y and H_2O . The dashed line indicates the boundary of the polar vortex.



8 p.p.b.v. in the Arctic (Fig. 2) is consistent with removal of the condensed phase as NAT.

The difficulty in attributing denitrification to the sedimentation of NAT or NAT-like aerosols is that the mass available for condensation limits the particle size if growth takes place on all available sites. Assuming a concentration of condensation nuclei in the lower polar stratosphere of 10 cm^{-3} at 15–20 km^{21,22}, an equivalent HNO_3 concentration of 10 p.p.b.v. creates type I particles characterized by a log-normal size distribution with a mode diameter near $0.5 \mu\text{m}$. The sedimentation velocity of a $0.5\text{-}\mu\text{m}$ particle is $<0.01 \text{ km per day}$ at 20 km (ref. 23). This rate is too low to account for the removal over the vertical scale noted in Fig. 3 if it is to occur in winter. The time required is difficult to determine from observations, but may be as low as a few days for H_2O (refs 4, 12). In addition, denitrification in the Arctic occurred over a few weeks at most, and perhaps occurred during a specific tropospheric forcing event²⁴.

This 'time difficulty' is resolved by allowing for selective nucleation and growth of pre-existing sulphate aerosols^{25,26}. Selective nucleation would limit growth to a small fraction of available sites, produce larger particles for the same mass deposition, and increase the sedimentation or denitrification rate, which is proportional to the square of the radius of a spherical particle. Observations of bimodal aerosol distributions in the Antarctic by Hofmann *et al.*²⁵ provide evidence for selective nucleation of large aerosols ($>0.6 \mu\text{m}$ diameter) where growth was limited to ~ 1 in 1,000 of available sites. The nucleation of sulphate aerosols under saturation conditions for NAT or ice is uncertain²⁷. The cooling rate near saturation is probably a factor, with large cooling rates resulting in the nucleation of all sites, whereas low to moderate rates possibly favour a few. The vortices clearly provide a wide range of cooling rates with measurements and meteorological analyses indicating values from $<1 \text{ K per day}$ to many K per day^{28,29}. Alternatively, ice

particles falling into a region of NAT saturation may simulate selective nucleation^{30,31}. The uptake of mass by these falling aerosols may be sufficient to prevent the occurrence of a large saturation ratio, necessary for the activation of a large fraction of the available particles.

Intense dehydration with denitrification in the Antarctic could evolve in two ways. First, denitrification with minimal dehydration could occur early in an air parcel history followed by more intense dehydration when temperatures fall below the frost point. Second, the removal of both species could occur simultaneously. As the vapour pressure of HNO_3 over ice is expected to be lower than that over NAT, HNO_3 could be incorporated into ice particles either directly, as they are growing at temperatures below the frost point, or in the form of NAT aerosols which serve as condensation nuclei for H_2O condensation^{27,32}. In general, denitrification may occur as a combination of both processes with the balance determined by the temperature history of air parcels as they approach and remain below the frost point.

The observed systematic differences in the denitrification and dehydration between the two polar data sets can be explained by temperature differences in the two regions. The minimum temperatures from the model analyses of the UK Meteorological Office as shown in Fig. 4 for the 30 hPa ($\sim 24 \text{ km}$) and 50 hPa ($\sim 21 \text{ km}$) levels provide a uniform comparison. In the Antarctic in 1987, temperatures remained below the 3.5-p.p.m.v. frost point at these pressure levels from June until early and mid-September, respectively⁴, consistent with the observed dehydration in this season. If 2 p.p.m.v. H_2O were incorporated into a particle distribution with a concentration of 10 cm^{-3} or lower, the duration of low temperatures is sufficient to allow sedimentation over several kilometres. In 1989 in the Arctic, the frost point for 5 p.p.m.v. was reached for only a few days in late January and early February. The limited Arctic dehydration observed is

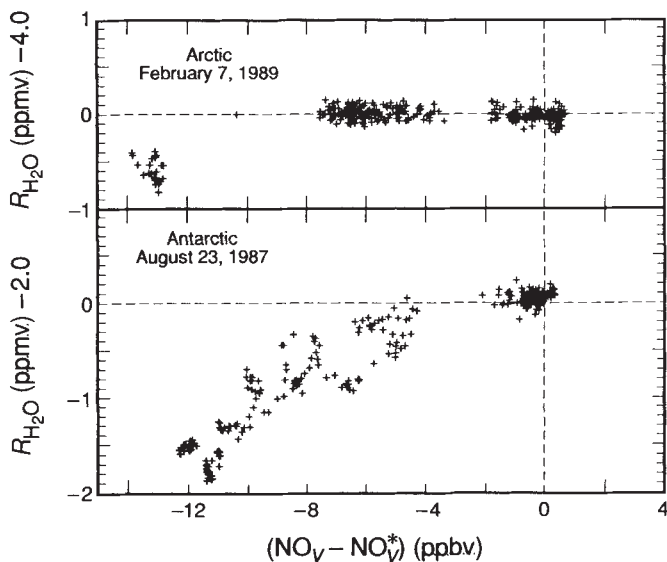


FIG. 2 Plot of the quantitative relationship of denitrification with dehydration for the flight legs in Fig. 1. Denitrification is quantified as $(\text{NO}_y - \text{NO}_y^*)$ with NO_y^* defined in the caption of Fig. 1. Dehydration is quantified by changes in residual water vapour, $R_{\text{H}_2\text{O}}$, from values at the vortex boundary. $R_{\text{H}_2\text{O}}$ is defined as measured H_2O minus H_2O produced from CH_4 (refs 4, 10):

$$R_{\text{H}_2\text{O}} = \text{H}_2\text{O} - 2[1.6 - \text{CH}_4] \quad (3)$$

where units are p.p.m.v. and the value 1.6 p.p.m.v. represents the tropopause or initial value of CH_4 in air entering the stratosphere. Continuous CH_4 values are derived from individual grab-sample measurements, and the average slope, $\Delta\text{N}_2\text{O}/\Delta\text{CH}_4$, in both data sets is $\sim 0.0037 \text{ p.p.b.v./p.p.m.v.}$. The values of 2 p.p.m.v. and 4 p.p.m.v. are the values of $R_{\text{H}_2\text{O}}$ for the Antarctic and Arctic, respectively, in unperturbed air near the vortex boundary in each data set. Some dehydration may be expected for the most intensely denitrified air parcels because the temperature difference between NAT and ice saturation decreases as NO_y decreases¹⁹.

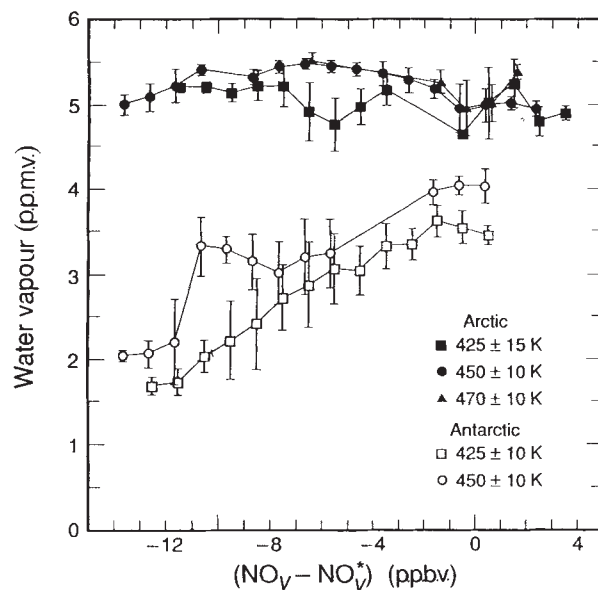


FIG. 3 Average scatter plot of H_2O and $(\text{NO}_y - \text{NO}_y^*)$ measurements. Open symbols represent measurements from seven Antarctic flights at different potential temperatures⁶. Solid symbols represent measurements from nine Arctic flights⁸. Intense denitrification was observed on all Antarctic flights, but only on three of the later Arctic flights. The data are 10-s average measurements sorted into intervals of 1 p.p.b.v. $(\text{NO}_y - \text{NO}_y^*)$ in each potential temperature interval. The vertical bars represent the sample standard deviation of a minimum of 10 points. The contribution of NO_y in aerosol particles is eliminated using the observed particle volume⁶. The simultaneous measurement of N_2O for each data point is $<160 \text{ p.p.b.v.}$ in the Antarctic and $<180 \text{ p.p.b.v.}$ in the Arctic, thereby restricting the observations to near and inside the vortex boundary in each data set. The N_2O limits are equivalent to a lower limit of unperturbed NO_y of ~ 10 and 8 p.p.b.v., respectively, as calculated from equation (2).

consistent with fewer occurrences of temperatures at or below the frost point. The relative differences between the 1987 Antarctic and 1989 Arctic temperatures are similar to the average differences for the period 1980–89, but temperatures are below the period average within both data sets³³.

The brief period of frost-point temperatures in the Arctic at 30 and 50 hPa is associated with upper-level cooling driven by a major tropospheric forcing event^{24,34}. The duration of the forcing means that a significant fraction of vortex air was exposed to temperatures at and below the frost-point temperatures. The principal ER-2 observations of denitrification are from flights in early February, several days after this event and nearly one month after the first occurrence of frost point at the 30-hPa level (14 January). In contrast, minimum temperatures were below NAT saturation for nearly two months at the 50-hPa level, which is closer to ER-2 flight levels. These relations in time suggest that perhaps temperatures near the frost point are required for intense denitrification, with or without accompanying intense dehydration. A similar conclusion is reached from a chemical modelling study along air parcel trajectories which coincide with the ER-2 flight tracks³⁴.

Figures 1 and 3 highlight the asymmetry in H₂O mixing ratios between the winter hemispheres, with lower values observed in the Antarctic at corresponding latitudes and potential temperatures. These differences are reflected in calculations of R_{H_2O} with considerably smaller values found in the Antarctic^{4,35}. Water vapour values in the tropics and mid-latitudes above 18 km do not show a comparable asymmetry³⁶. The low Antarctic values have been attributed to the transport of dehydrated air from the vortex interior to lower latitudes^{4,12,35}. In contrast, similar differences are not found in the NO_y/N₂O relation

between the hemispheres⁹. To the extent to which the NO_y/N₂O relation defines a reference state, the Antarctic H₂O data provide evidence of intense dehydration not accompanied by comparable denitrification. The uncertainties in the microphysical processes of removal allow for the possibility of dehydration without denitrification. For example, if an air parcel temperature falls below the NAT frost point with a cooling rate large enough to activate all available nuclei, a large fraction of available HNO₃ may be sequestered on non-sedimenting (<0.01 km per day) NAT particles. With continued cooling below the frost point, ice may selectively nucleate on a fraction of the existing NAT particles. These particles subsequently grow through ice deposition under conditions of low HNO₃ vapour, and sediment in a time period in which transfer of HNO₃ vapour between particles cannot be effective^{32,37}. □

Received 27 December 1989; accepted 22 February 1990.

- Brune, W. H., Toohey, D. W., Anderson, J. G. & Chan, K. R. *Geophys. Res. Lett.* (submitted).
- Tuck, A. F., Watson, R. T., Condon, E. P., Margitan, J. J. & Toon, O. B. *J. geophys. Res.* **94**, 11687–11737 (1989).
- Fahey, D. W. et al. *J. geophys. Res.* **94**, 11299–11315 (1989).
- Kelly, K. K. et al. *J. geophys. Res.* **94**, 11317–11357 (1989).
- Podolske, J. R., Loewenstein, M., Strahan, S. E. & Chan, K. R. *J. geophys. Res.* **94**, 16767–16772 (1989).
- Fahey, D. W. et al. *J. geophys. Res.* **94**, 16665–16681 (1989).
- Proffitt, M. H. et al. *J. geophys. Res.* **94**, 11437–11448 (1989).
- Kawa, S. R., Fahey, D. W., Anderson, L. C., Loewenstein, M. & Chan, K. R. *Geophys. Res. Lett.* (in the press).
- Fahey, D. W. et al. *Nature* (submitted).
- Jones, R. L. et al. *Q. J. R. met. Soc.* **112**, 1127–1143 (1986).
- Watterson, I. G. & Tuck, A. F. *J. geophys. Res.* **94**, 16511–16525 (1989).
- Tuck, A. F. *J. geophys. Res.* **94**, 11687–11737 (1989).
- Toon, O. B., Hamill, P., Turco, R. P. & Pinto, J. *Geophys. Res. Lett.* **13**, 1284–1287 (1986).
- Crutzen, P. J. & Arnold, F. *Nature* **324**, 651–655 (1986).
- McElroy, M. B., Salawitch, R. J. & Wofsy, S. C. *Geophys. Res. Lett.* **13**, 1296–1299 (1986).
- Pueschel, R. F. et al. *J. geophys. Res.* **94**, 11271–11284 (1989).
- Arnold, F., Schlager, H., Hoffmann, J., Metzinger, P. & Spreng, S. *Nature* **342**, 493–497 (1989).
- Poole, L. R. & McCormick, M. P. *J. geophys. Res.* **93**, 8423–8430 (1988).
- Hanson, D. & Mauersberger, K. *Geophys. Res. Lett.* **15**, 855–858 (1988).
- Kondo, Y. et al. *Geophys. Res. Lett.* (in the press).
- Wilson, J. C. et al. *J. geophys. Res.* **94**, 16437–16448 (1989).
- Hofmann, D. J. *Rev. Geophys.* **26**, 113–130 (1990).
- Kasten, F. J. *J. appl. Met.* **7**, 944–947 (1968).
- McCormick, M. P. et al. *Geophys. Res. Lett.* (submitted).
- Hofmann, D. J., Rosen, J. M., Harder, J. W. & Hereford, J. V. *J. geophys. Res.* **94**, 11253–11269 (1989).
- Salawitch, R. J., Gobbi, G. P., Wofsy, S. C. & McElroy, M. B. *Nature* **339**, 525–527 (1989).
- Toon, O. B., Turco, R. P., Jordan, J., Goodman, J. & Ferry, G. *J. geophys. Res.* **94**, 11359–11380 (1989).
- McKenna, D. S. et al. *J. geophys. Res.* **94**, 11641–11668 (1989).
- Gary, B. L. et al. *Geophys. Res. Lett.* (submitted).
- Wofsy, S. C. et al. *Geophys. Res. Lett.* (submitted).
- Dye, J. E. et al. *Geophys. Res. Lett.* (submitted).
- Turco, R. P., Toon, O. B. & Hamill, P. *J. geophys. Res.* **94**, 16493–16510 (1989).
- Scientific Assessment of Stratospheric Ozone: 1989 Rep. 20, Ch. 1 (World Meteorological Organization, 1990).
- Jones, R. L., McKenna, D. S., Poole, L. R. & Solomon, S. *Geophys. Res. Lett.* (submitted).
- Kelly, K. K. et al. *Geophys. Res. Lett.* (in the press).
- Hansen, A. R. & Robinson, G. D. *J. geophys. Res.* **94**, 8474–8484 (1989).
- Toon, O. B., Turco, R. P. & Hamill, P. *Geophys. Res. Lett.* (submitted).

ACKNOWLEDGEMENTS. We thank B. W. Gandrud, D. J. Hofmann, L. R. Poole, S. Solomon, O. B. Toon and S. C. Wofsy. This research is supported in part by NASA as part of the AAQE and AASE aircraft missions. Travel costs for NOAA personnel were provided by the Chemical Manufacturer's Association for AAQE, and by NOAA Headquarters for AASE.

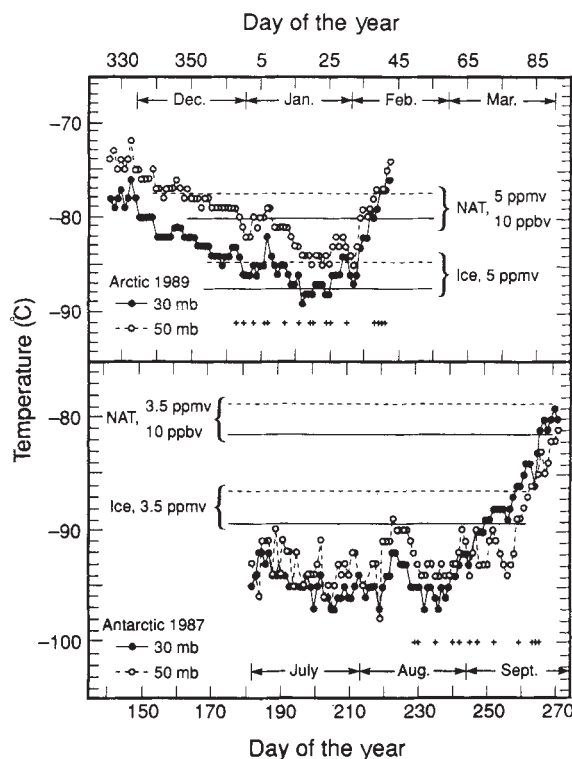


FIG. 4 Plot of the minimum temperatures poleward of 60° latitude at the 30-hPa and 50-hPa (30 and 50 mbar) levels in the UK Meteorological Office analyses for the Antarctic and Arctic missions. The crosses in each panel represent flight days of the ER-2 aircraft for each mission. The data sets are offset horizontally so that equivalent seasons are vertically aligned. The condensation temperatures for NAT and ice are indicated for each level for nominal values of HNO₃ and/or H₂O mixing ratios. Occasionally, during periods of rapidly changing conditions or intense disturbances, the analysed temperatures are found to be in error by a few degrees or more when compared with observations. An example is the period near ice saturation in the Arctic, during which observations showed lower temperatures.

Natural variability of the climate system and detection of the greenhouse effect

T. M. L. Wigley & S. C. B. Raper

Climatic Research Unit, University of East Anglia, Norwich NR4 7TJ, UK

GLOBAL mean temperatures show considerable variability on all timescales. The causes of this variability are usually classified as external or internal¹, and the variations themselves may be usefully subdivided into low-frequency variability (timescale ≥ 10 years) and high-frequency variability (≤ 10 years). Virtually nothing is known about the nature or magnitude of internally generated, low-frequency variability. There is some evidence from models, however, that this variability may be quite large^{1,2}, possibly causing

fluctuations in global mean temperature of up to 0.4°C over periods of thirty years or more (see ref. 2, Fig. 1). Here we show how the ocean may produce low-frequency climate variability by passive modulation of natural forcing, to produce substantial trends in global mean temperature on the century timescale. Simulations with a simple climate model are used to determine the main controls on internally generated low-frequency variability, and show that natural trends of up to 0.3°C may occur over intervals of up to 100 years. Although the magnitude of such trends is unexpectedly large, it is insufficient to explain the observed global warming during the twentieth century.

Observed global mean temperatures show considerable inter-annual variability. If the data of Jones *et al.*³ are filtered using a 10-year gaussian filter, overall, the high-frequency and the low-frequency standard deviations over 1867–1982 are 0.175°C , 0.078°C and 0.148°C , respectively. Thirty-three per cent of the high-frequency variance can be identified with the El Niño/Southern Oscillation phenomenon (ENSO). If this is factored out by simple linear regression against the Darwin–Tahiti pressure index of the Southern Oscillation⁴, the remaining high-frequency standard deviation is 0.063°C , a value that is remarkably stable in time: 1867–88, 0.069°C ; 1889–1910, 0.062°C ; 1911–32, 0.061°C ; 1933–54, 0.064°C ; 1955–76, 0.065°C . A very small fraction of this residual variability is due to the effects of major volcanic eruptions which have an identifiable cooling signal lasting for a few years^{5–7}. The remainder is probably due largely to the internal variability of the atmosphere, manifest at least partly as global-scale changes in the cloud radiation budget⁸. If these variations are the high-frequency component of random, white-noise forcing, they must, because of the thermal inertia of the oceans, produce global mean temperature changes that have enhanced low-frequency variability. We aim to determine the magnitude of this low-frequency variability.

Global mean temperature changes ($\Delta T(t)$) in response to forcing $\Delta Q(t)$ can be modelled using an energy-balance model of the form^{9,10}

$$\phi\rho ch \frac{d\Delta T}{dt} + \lambda\Delta T = \Delta Q - \Delta F \quad (1)$$

$$\Delta F = -\phi\rho c \left(K \frac{\partial\Delta\theta}{\partial z} + W\Delta\theta \right)_{z=0} \quad (2)$$

$$\frac{\partial\Delta\theta}{\partial t} - W \frac{\partial\Delta\theta}{\partial z} = K \frac{\partial^2\Delta\theta}{\partial z^2} \quad (3)$$

In equation (1), ϕ (≈ 0.7) is a factor accounting for land/sea differences in heat capacity¹⁰, ρ is the density and c the specific heat capacity of ocean water, h is the mean depth of the ocean mixed layer, λ is a feedback parameter¹¹ and ΔF is the flux of heat out of the mixed layer into the deeper ocean. ΔF (equation (2)) is estimated using an upwelling diffusive parameterization of vertical mixing. $\Delta\theta(z, t)$ is the temperature change at depth z below the mixed layer, W is the upwelling velocity and K is the vertical diffusivity.

If ΔQ is white-noise forcing with a uniform spectrum, then the thermal inertia of the ocean will produce a redder spectrum with a preponderance of variance at low frequencies, as pointed out by Hasselmann¹². Although the longer-term climate consequences have been explored in the frequency domain¹³, they have never been properly quantified in the context of climate change on a 10–100-yr timescale.

Equations (1)–(3) can be solved exactly in the frequency domain, using the boundary conditions $\Delta\theta_{z=0} = \Delta T$ and $\partial\theta/\partial z$ tending to zero for large z , to give

$$S_T = S_Q \lambda^{-1} \left\{ \left(1 - \frac{\tau W}{2h} + \frac{\tau}{2h} \sqrt{\frac{v^2 + W^2}{2}} \right)^2 + \left(2\pi f \tau + \frac{\tau}{2h} \sqrt{\frac{v^2 - W^2}{2}} \right)^2 \right\}^{-1} \quad (4)$$

where $S_Q(f)$ is the power spectrum of the forcing at frequency

f (yr^{-1}), $S_T(f)$ is the power spectrum of the mixed layer temperature, τ is a mixed-layer timescale

$$\tau = \phi\rho ch / \lambda \quad (5)$$

($\tau \approx 2\text{--}8$ yr) and v is a characteristic velocity defined by

$$v^4 = W^4 + 64\pi^2 f^2 K^2 \quad (6)$$

The ratio $\sqrt{S_T}/\sqrt{S_Q}$ may be considered as a frequency-dependent climate sensitivity. Climate sensitivity is normally defined as the inverse of the feedback parameter, to which $\sqrt{S_T}/\sqrt{S_Q}$ tends at low frequencies.

At high frequencies ($f \geq 10^{-1} \text{ yr}^{-1}$),

$$S_T = S_Q / (2\pi\phi\rho ch f)^2 \quad (7)$$

The high-frequency response is therefore independent of the feedback parameter (and hence the climate sensitivity). This means that the response of the climate system to high-frequency forcings such as volcanic eruptions and the seasonal insolation cycle must be virtually independent of the sensitivity. High-frequency information is therefore of little value in trying to estimate, empirically, the climate sensitivity. This is an obvious, but little appreciated result.

At low frequencies ($f \leq 10^{-3} \text{ yr}^{-1}$), S_T is approximately constant:

$$S_T = S_Q / \lambda^2 \quad (8)$$

This follows from the fact that surface temperature fluctuations track equilibrium values when the forcing period is much greater than the ocean response time. As $S_Q \approx (1 \text{ W m}^{-2})^2$ (see below), $S_T(f=0)$ varies between $(0.3^\circ\text{C})^2$ at $\lambda = 3$ and $(1.0^\circ\text{C})^2$ at $\lambda = 1$. Thus, passive internal variability can account for only a small fraction of climate variability on the ice-age timescale, in agreement with current thinking about the importance of orbital forcing. Qualitatively similar results to those above have been given by Lemke¹³, but the variability derived here is much lower because we have used more realistic feedback parameter values.

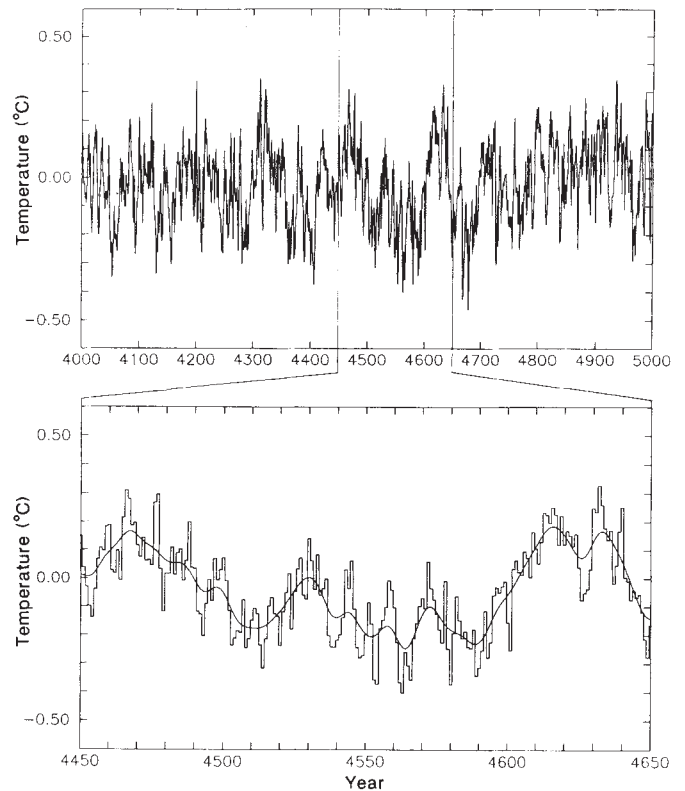


FIG. 1 Simulation of internally generated natural variability of global mean temperature ($\Delta T_{2x} = 4.0^\circ\text{C}$, $K = 1.0 \text{ cm}^2 \text{ s}^{-1}$). The upper panel is the fifth 1,000 years of a 100,000-yr run, with 200 years enlarged in the lower panel. The smooth curve in the lower panel shows filtered values to highlight the lower frequency variations.

TABLE 1 Statistical characteristics of model-generated natural variability for different model parameters

	1.5				3.0				4.5			
Sensitivity, ΔT_{2x}	0.0*	0.5	1.0	2.0	0.0*	0.5	1.0	2.0	0.0*	0.5	1.0	2.0
Diffusivity, $K =$	0.73	0.64	0.61	0.59	0.83	0.72	0.69	0.65	0.87	0.75	0.72	0.68
Autocorrelation (r_1)												
Standard deviations†												
σ_1	0.14	0.12	0.12	0.12	0.19	0.15	0.14	0.14	0.23	0.16	0.16	0.15
σ_{10}	0.09	0.08	0.08	0.08	0.14	0.11	0.11	0.10	0.19	0.13	0.12	0.12
Trends (90% band)‡												
10 yr	±0.39	±0.31	±0.29	±0.28	±0.49	±0.35	±0.33	±0.31	±0.55	±0.37	±0.34	±0.31
30 yr	±0.30	±0.24	±0.24	±0.24	±0.46	±0.32	±0.30	±0.29	±0.59	±0.36	±0.33	±0.31
100 yr	±0.18	±0.17	±0.17	±0.18	±0.31	±0.25	±0.25	±0.25	±0.43	±0.30	±0.29	±0.29

Mixed-layer depth was taken as 70 m and upwelling rate chosen to ensure that W/K remained constant, with $W = 4 \text{ m yr}^{-1}$ for $K = 1 \text{ cm}^2 \text{ s}^{-1}$. Low-frequency variability is larger for larger mixed-layer depth. The other model parameter that influences results is the temperature of the sinking water, which we have taken to be constant here (that is, $\Pi = \Delta T_a / \Delta T = 0$; see ref. 24 for support of this assumption). Taking $\Pi > 0$ reduces the low-frequency variability slightly. Lag-1 autocorrelations are the means of values for 100-year samples. Model runs were 100,000 years. Temperatures in $^{\circ}\text{C}$.

* W set to zero too (corresponding to a mixed-layer ocean model).

† σ_1 is the standard deviation of the full time series of annual values. σ_{10} is the standard deviation of non-overlapping 10-year means.

‡ 90% of all values lay within the limits shown.

In the time domain, important results can be obtained analytically^{10,14–16}, but in the context of stochastic forcing it is more informative to proceed through numerical integration of the equations. We have used a generalization of this type of model in which the Earth is divided into land and ocean 'boxes' in each hemisphere (see ref. 17). To facilitate comparison with greenhouse studies, we specify climate sensitivity in terms of the equilibrium warming for a doubling of atmosphere CO_2 , $\Delta T_{2x} = 1.5\text{--}4.5^{\circ}\text{C}$, related to λ by

$$\lambda^{-1} = \Delta T_{2x} / \Delta Q_{2x} \quad (9)$$

where ΔQ_{2x} is taken as 4.39 W m^{-2} (refs 18, 19). Vertical heat transport in the ocean is assumed to be an upwelling diffusive process.

Our aim is to determine the magnitude of the decadal to century timescale variations in global mean temperature under the constraint that they are compatible with the observed high-frequency variations. The model is therefore forced with random interannual variations in ΔQ ; specifically, global mean gaussian white-noise forcing.

The total variance of the forcing required to match observed high-frequency temperature variations (σ_Q^2) can be estimated from the spectral solution, knowing the high-frequency variance of the global mean temperature data. The implied value of σ_Q is $\sim 1 \text{ W m}^{-2}$. This value is consistent with the observed month-to-month variations in the Earth's radiation budget due to cloudiness variations⁸, which have a standard deviation of 3.3 W m^{-2} . If month-to-month variations are uncorrelated, then the interannual standard deviation should be $\sim 1 \text{ W m}^{-2}$. This may be interpreted as a forcing term provided that the radiation budget changes are not themselves dependent on temperature changes (M. I. Hoffert, personal communication). Although such variations may, in general, be affected by ENSO events, the data in ref. 8 come from months with relatively small values of the Southern Oscillation Index (SOI) and show no correlation with SOI.

After forcing the model with gaussian white noise with $\sigma_Q = 1 \text{ W m}^{-2}$, we then adjusted the forcing (by a factor of 0.72 to 1.04 depending on model parameters) so as to match precisely the observed high-frequency variations in ΔT ($\sigma_T = 0.063^{\circ}\text{C}$), and then examined the low-frequency characteristics of the output. The σ_T value used here corresponds to the global mean temperature record with ENSO factored out. This is appropriate, primarily because the spectrum of the ENSO part of the signal (as determined by the Darwin–Tahiti pressure index used here²⁰) has negligible low-frequency variance and so cannot lead to any noticeable low-frequency response. Simulations of 100,000 years length were performed for a wide range of values of the parameters that control model output.

The low-frequency character of the simulated global mean temperature variations that is of particular interest here is that around the century timescale, as this is particularly pertinent to

the issue of greenhouse-gas forcing. This is in the frequency region in which the results begin to show appreciable λ dependence (see equation (4)). To compare the simulations with observations most readily, we examined the distribution of trends (dT_L) of durations (L) from 10 to 1,000 years.

Typical simulated variations in global mean temperature are shown in Fig. 1. These variations are qualitatively similar to observed changes over the past 100 years or more³. The 10–100-yr timescale fluctuations can be seen to be substantial. These are the result of the thermal inertia of the oceans acting in a purely passive mode, independent of any external forcing changes and independent of active oceanic effects which might be associated with circulation changes.

The magnitude of this passive internal variability depends on the model parameters, as shown in Table 1. (The results shown in Table 1 are for globally coherent forcing, but they are insensitive to the degree of aggregation of the forcing.) Low-frequency variability, as judged by any of the parameters shown, increases with increasing ΔT_{2x} and decreasing K as predicted by the analytical results. Simulated lag-1 autocorrelations (r_1 ; that is, the correlation between successive annual mean values) are similar to those observed in the data of Jones *et al.*³. The data, however, behave quite differently from what would be expected if a first-order autoregressive process described the temperature variations (as assumed, for example, in ref. 21), pointing to the importance of higher-order effects (see ref. 22). The 100-year trend confidence limits, although substantial, are much less than the observed trend of 0.5°C over the past 100 years. In the sense that passive internal variability represents the most basic underlying natural variability of the climate system, this result indicates that the observed warming trend is highly statistically significant. Based on the results presented here, the 90% confidence limits for a trend of 0.5°C range from $0.34\text{--}0.66^{\circ}\text{C}$ to $0.24\text{--}0.76^{\circ}\text{C}$ depending on climate sensitivity. (Confidence limits of 90% are given here because the significance test for a warming trend is only one-tailed.)

Table 1 also shows that the variability for a mixed-layer-only ocean ($K = 0$, $W = 0$) is appreciably greater than for non-zero K (as can also be seen from equation (4)). Simulations by Robock¹ and Hansen *et al.*², which are based on mixed-layer ocean formulations, have therefore probably overestimated the degree of low-frequency variability that might arise through this internal mechanism.

These results have direct bearing on the value of the climate sensitivity parameter (that is, on ΔT_{2x} or λ), and hence on the issue of detecting the greenhouse effect on climate. Gilliland and Schneider²³ and Wigley and Raper¹⁷ have noted that the observed warming, if interpreted solely as a greenhouse signal, implies a rather low value for the climate sensitivity, $\Delta T_{2x} = 1.3\text{--}2.0^{\circ}\text{C}$ (the range of values arises from uncertainties in other model parameters). If one allows for passive internal variability, the range of possible ΔT_{2x} values increases substantially, to

1.0–2.8 °C. Higher values can be admitted only if the present model substantially underestimates the lag effect of oceanic thermal inertia, or if other external forcing factors or internal effects (such as ocean circulation changes) are invoked. Because of a dearth of suitable data, however, these possibilities cannot be ruled out, so neither can one rule out high values of climate sensitivity. Thus, although we still cannot claim to have unequivocally detected the greenhouse effect, our consideration of passive internal variability virtually eliminates this internal mechanism as a possible sole explanation of the recent global warming. Although our analysis has increased the range of uncertainty in the empirical value of ΔT_{2x} , this has served mainly to bring the empirical result more in line with the range of ΔT_{2x} values estimated from general circulation models (namely, 1.5–4.5 °C), strengthening the degree of consistency between model predictions and observations. □

Received 4 January; accepted 9 February 1990.

1. Robock, A. *J. Atmos. Sci.* **35**, 1111–1122 (1978).
2. Hansen, J. et al. *J. Geophys. Res.* **93**, 9341–9364 (1988).
3. Jones, P. D., Wigley, T. M. L. & Wright, P. B. *Nature* **322**, 430–434 (1986).
4. Jones, P. D. *Clim. Mon.* **17**, 80–89 (1989).
5. Kelly, P. M. & Sear, C. B. *Nature* **311**, 740–743 (1984).
6. Sear, C. B., Kelly, P. M., Jones, P. D. & Goodess, C. N. *Nature* **330**, 365–367 (1987).
7. Bradley, R. S. *Clim. Change* **12**, 221–243 (1988).
8. Ramanathan, V. et al. *Science* **243**, 57–63 (1989).
9. Hoffert, M. I. & Flannery, B. P. in *Projecting the Climatic Effects of Increasing Carbon Dioxide* (eds MacCracken, M. C. & Luther, F. M.) 149–190 (US Department of Energy, Washington, DC, 1985).
10. Wigley, T. M. L. & Schlesinger, M. E. *Nature* **315**, 649–652 (1985).
11. Schlesinger, M. E. *Clim. Dynam.* **1**, 35–51 (1986).
12. Hasselmann, K. *Tellus* **28**, 473–485 (1976).
13. Lemke, P. *Tellus* **29**, 385–392 (1977).
14. Watts, R. G. *J. Geophys. Res.* **90**, 8067–8070 (1985).
15. Lebedeff, S. A. *J. Geophys. Res.* **93**, 14243–14255 (1988).
16. Morantini, M. & Watts, R. G. *J. Geophys. Res.* (in the press).
17. Wigley, T. M. L. & Raper, S. C. B. *Nature* **330**, 127–131 (1987).
18. Kiehl, J. T. & Dickinson, R. E. *J. Geophys. Res.* **92**, 2991–2998 (1987).
19. Wigley, T. M. L. *Geophys. Res. Lett.* **14**, 1135–1138 (1987).
20. Ropelewski, C. F. & Jones, P. D. *Mon. Weath. Rev.* **115**, 2161–2165 (1987).
21. Wigley, T. M. L. & Jones, P. D. *Nature* **292**, 205–208 (1981).
22. Tsonis, A. A. & Elsner, J. B. *Geophys. Res. Lett.* **16**, 795–797 (1989).
23. Gilliland, R. L. & Schneider, S. H. *Nature* **310**, 38–41 (1984).
24. Harvey, L. D. D. & Schneider, S. H. *J. Geophys. Res.* **90**, 2191–2205 (1985).

ACKNOWLEDGEMENTS. This work was supported by a grant from the US Department of Energy, Carbon Dioxide Research Division.

Deep structure of an Archaean greenstone terrane

A. G. Green^{*†}, B. Milkereit[‡], L. J. Mayrand[‡],
J. N. Ludden[‡], C. Hubert[‡], S. L. Jackson[§],
R. H. Sutcliffe[§], G. F. West^{||}, P. Verpaal[¶]
& A. Simard[¶]

^{*}GRANSIR, 10 Chemin de la Raye, 1024 Ecublens, Switzerland

[†]Geological Survey of Canada, Ottawa, Ontario K1A 0Y3, Canada

[‡]Université de Montréal, Département de Géologie, Montréal, Québec H3C 3J7, Canada

[§]Ontario Geological Survey, 77 Grenville Street, Toronto, Ontario M7A 1W4, Canada

^{||}University of Toronto, Physics Department, Toronto, Ontario M5S 1A7, Canada

[¶]Ministère de l'Énergie et des Ressources du Québec, Val d'Or, Québec J9P 3L4, Canada

MORE than half of the continental crust was formed before about 2,500 million years ago¹, yet there is still considerable uncertainty regarding the structure and evolution of the ancient Archaean cratons^{2–7}. Modern analogues proposed for the setting of Archaean greenstone belts include aborted intra-cratonic rifts, continental margins and various oceanic settings^{2–7}; alternatively, Archaean tectonics may have been quite different from the plate tectonics we see today^{2–4,6}. The Abitibi belt in Canada is the largest Archaean greenstone terrane³, and has been a touchstone for many models because its surface geology is relatively well known^{8–10}, it

contains significant gold and base metal deposits¹⁰, and it is less deformed and has suffered less erosion than most Archaean terranes. Here we present deep seismic reflection data from part of the Abitibi belt, which reveal extensive subhorizontal and gently dipping reflection zones, truncated by steeply inclined discontinuities which project to major fault zones at the surface. Our preliminary model of the Abitibi belt comprises a mostly juvenile magmatic crust, similar to that exposed in the nearby Kapuskasing uplift^{11,12}, dissected and imbricated by deep-penetrating discontinuities associated with thrust and strike-slip faulting. In view of the similarity of greenstone belts worldwide^{2–7,13}, our seismic sections provide new constraints on the general processes that have affected Archaean metavolcanic terranes.

Moderate to steeply dipping komatiitic, tholeiitic and bimodal (mafic/felsic) metavolcanic rocks dominate the surveyed area of the Abitibi greenstone belt (AGB; Fig. 1)^{8,9}. Partly bounding the youngest and largest bimodal sequence are the Larder Lake/Cadillac (LCF) and Porcupine/Destor (PDF) fault zones. These two huge east-trending fault zones, which extend for hundreds of kilometres, contain many of Canada's rich gold deposits¹⁰. In the eastern portion of the study area the LCF juxtaposes fundamentally different terranes (Fig. 1), the metavolcanic AGB and the metasedimentary–plutonic Pontiac subprovince.

The two regional fault zones and intervening structures were seismically surveyed along three north–south corridors (Fig. 1). As previous seismic reflection surveying across North American greenstone terranes had only limited success in imaging shallow and deep structures^{14,15}, we collected data using both regional and high-resolution parameters (see Fig. 2 legend). Comparison of the high-resolution section 14B with the coincident regional section 14 (Fig. 2a, b) attests to the superior results determined in high-resolution mode. To obtain quality images of the ancient crystalline crust, especially at shallow depths, our testing demonstrated that it was necessary to have high input frequencies (20–80 Hz), close source and receiver spacing (<50 m), short source–receiver offsets (<3–4 km), and high density of coverage (more than sixtyfold).

The reflection patterns in the AGB data vary significantly according to the geology crossed (Fig. 3). They change from well defined regional reflections underlying metavolcanic rocks along all corridors to zones almost devoid of reflections beneath intrusions along portions of the line 14. Continuous layering below the Pontiac subprovince on line 14B differs markedly from the style of reflections in the region of AGB adjacent to the LCF (Fig. 2a), compatible with the observed abrupt changes in petrological, metamorphic and structural trends across the subprovince boundary. That no single pattern of reflectivity represents the study area is evidence that Archaean crustal sections are as structurally diverse as their Proterozoic and Phanerozoic counterparts (see also ref. 12).

The AGB seismic sections reveal prominent subhorizontal and shallow dipping reflection zones extending for 10–40 km in the interval of 1–7 s (corresponding to depths of 3–22 km using velocities from nearby refraction lines¹⁶); examples include E1 to E3 and F on line 12 (Figs 2c and 3), A to F on line 12A, and H, L, O and P on lines 14 and 14B (Figs 2a, b and 3). Some subhorizontal reflections occur at relatively shallow depths (6 km) beneath rock assemblages that have moderate to steeply dipping contacts and foliations at the surface^{8,9}. Beneath line 12, less prominent but 'homogeneous' reflectivity occurs in the lower crust between 5 s (16 km) and 11–11.5 s (35–38 km; Fig. 2c).

Many reflections seem to be truncated by steep discontinuities that project to surface traces of the major fault zones (Figs 2 and 3). The distinctive steps of the E1 to E3 reflections on section 12 underlie an area dissected by splays of the PDF, and to the east, below line 14, truncations of reflections and lateral changes of reflection dip across the PDF (subhorizontal to north dips south of the fault, and south dips to the north) allow the

nearly vertical fault zone to be mapped to at least 5 s. Along the southern margin of the AGB, close to line 14, the LCF has been observed dipping steeply to the north in a gold mine¹⁷. Consequently, on Fig. 3 we show a steep-dipping LCF merging with the curved band of events *N*, which separates the strongly reflective crustal section that is characteristic of the Pontiac subprovince from a narrow reflection-free zone immediately north of LCF (Fig. 2a). A steep discontinuity is also inferred below splays of the LCF on line 12A. Here, depth projections of the faults are outlined by truncations of reflections A to C and by contrasting reflection patterns (north dips south of the faults and subhorizontal to south dips to the north). Original field records and unmigrated stacked data demonstrate that genuine variations in seismic properties of rocks at depth, rather than changes in noise or near-surface conditions, best account for the lateral changes in reflectivity observed on the migrated seismic sections (Figs 2 and 3).

To interpret the origin of AGB subhorizontal reflections and the lithologies of intervening rocks we take advantage of a crustal model based on the nearby Kapuskasing uplift^{11,12}, where the upper two-thirds of AGB crust may be exposed. According to this model, the 6–13-km-thick layer above reflection zone E (lines 12 and 12A) comprises greenstone rocks, the middle layer bounded by E and F is composed of intraplated tonalites, and the lower layer beneath F contains layered felsic/mafic granulites and large lenses of anorthositic rock with the proportion of mafic and anorthositic rock inferred to increase with depth. Reflection zones E and F originate at appropriate depths for the major lithological/structural boundaries defined in the

Kapuskasing region, and their gently rolling topography is comparable to the 20–25-km-wide domal structures of Kapuskasing tonalites. Furthermore, prominent reflections from AGB lithological/structural boundaries are consistent with velocity/density measurements on appropriate Kapuskasing rocks¹⁸ and refraction velocities and wide-angle reflections recorded in western AGB¹⁶. Finally, the pattern of lower crustal reflectivity (below F, Fig. 2c) is similar to that recorded at shallow depths beneath exposed layered granulitic and anorthositic rocks of the Kapuskasing uplift¹².

Independent of the Kapuskasing model, the transition between crust and mantle is assigned to the distinct decrease in density of reflections at 11.0–11.5 s or a depth of 35–38 km (Figs 2c and 3), which compares with refraction Moho depths of 38–45 km and 35–38 km in the western and eastern AGB respectively^{16,19}.

The offsets and truncations of the subhorizontal reflections suggest that the two major discontinuities and their splays were generated by crustal-scale thrust and/or strike-slip faulting. In general, the fault zones dip steeply to at least 3–5 s (9–16 km), but it is not known whether they become listric at greater depth as shown in our preferred model (Fig. 3), or are truncated by younger features (for example, late intrusions). That reflections cannot be correlated across the LCF and its splays suggests an important component of lateral movement in addition to any normal or reverse displacements, consistent with independent structural evidence for strike-slip faulting^{10,20}. Except for reflection zone A below line 12A (Fig. 3), which may project to a north-dipping deformation zone and a gabbroic sill, none of

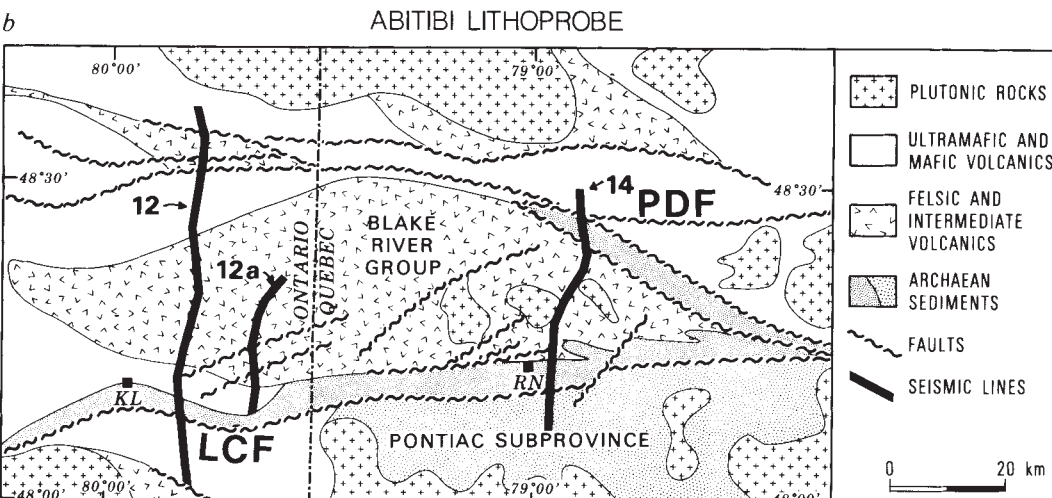
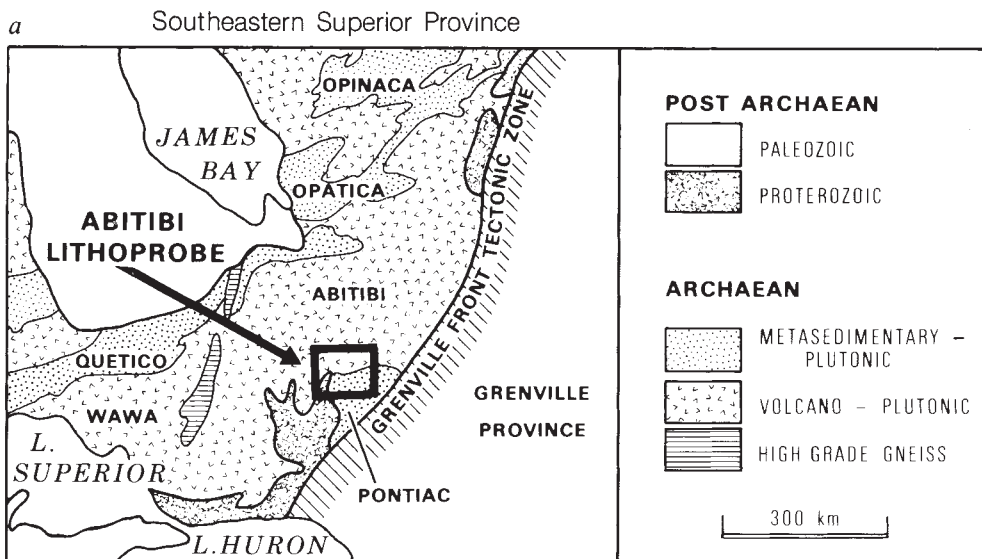


FIG. 1 a, Location of Abitibi LITHOPROBE project within the southern Canadian Shield; note the critical position of the survey area relative to the boundary between the Abitibi greenstone belt and the Pontiac subprovince. The area of high-grade rocks between the Abitibi and Wawa greenstone belts is the Kapuskasing uplift^{11,12}, on which the deep lithological/structural model is based. b, Simplified geology of southern Abitibi greenstone belt crossed by LITHOPROBE seismic reflection profiles 12, 12A and 14. LCF, Larder Lake/Cadillac fault zone; PDF, Porcupine/Destor fault zone; probable splays of the LCF crossed by line 12A are the Mulven Lake (MLF) and Misema Lake/Mist Lake (MMF) faults (see Fig. 3); RN, Rouyn/Noranda; KL, Kirkland Lake.

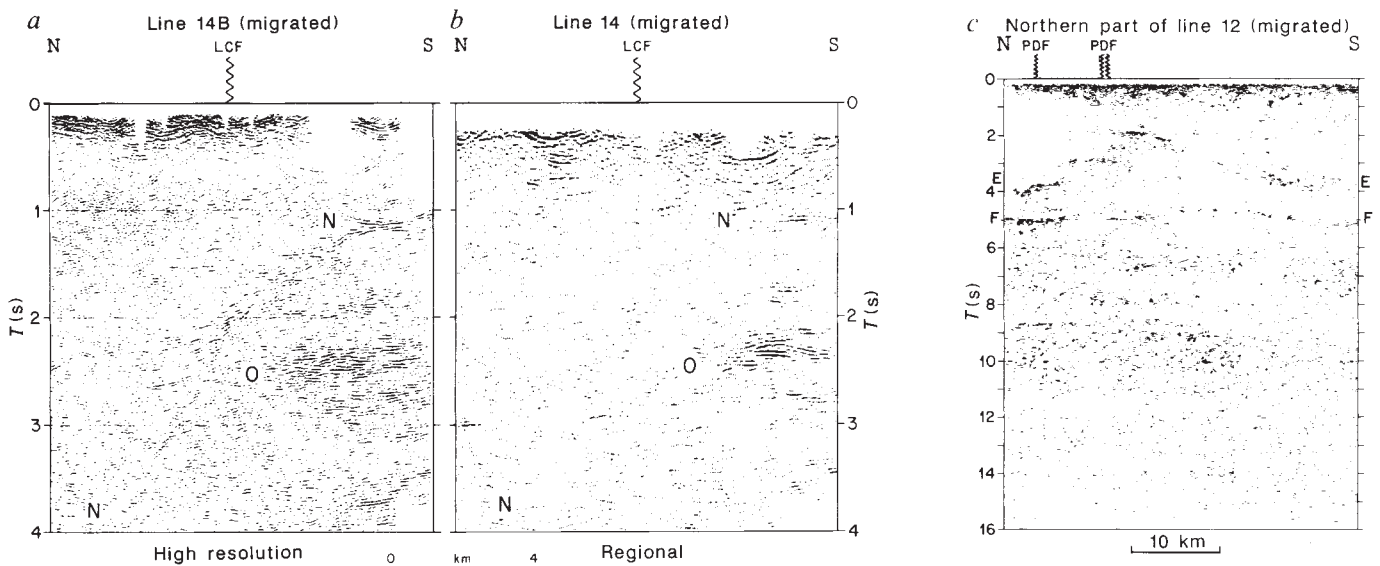


FIG. 2 Migrated seismic sections from *a*, south part of high-resolution line 14B, *b*, equivalent segment of regional line 14 and *c*, north half of regional line 12. (See geological map of Fig. 1 for location of seismic lines.) Note that events N-N are imaged on *a* but not *b*, illustrating the important role of acquisition parameters for imaging steep-dipping structures. Lines 12 and 14 were recorded with the following regional parameters: 12–52-Hz sweeps, 4-ms sampling interval, 4 vibrators, 50-m station spacing, 100-m

source interval, 240 channels, 60-fold coverage, 16-s records. Lines 12A, 14A and 14B were recorded with the following high-resolution parameters: 20–120 Hz sweeps, 2-ms sampling interval, 2 vibrators, 20-m station spacing, 20-m source interval, 240 channels, 120-fold coverage, 8-s correlated records. Migration velocities vary linearly from 5.75 km s^{-1} at the top to 6.25 km s^{-1} at the bottom of sections, which are plotted at $\sim 1:1$ for a velocity of 6 km s^{-1} .

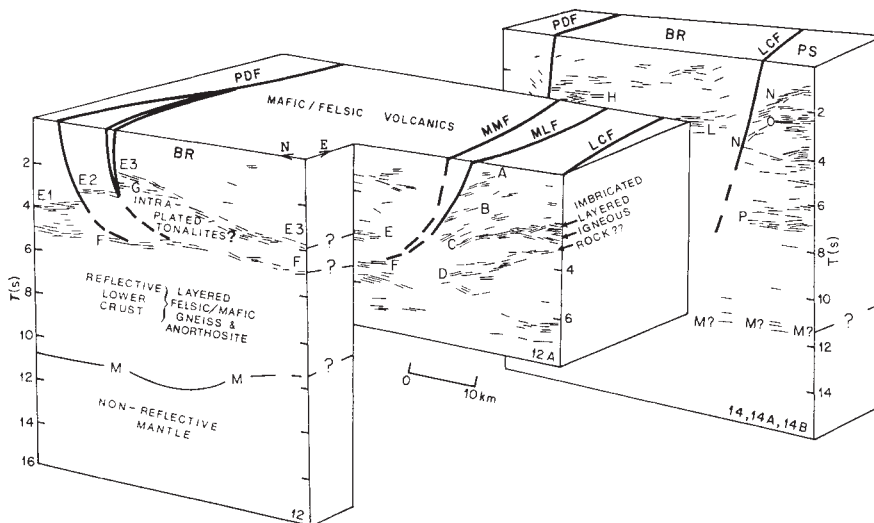


FIG. 3 Sketch of main features revealed by migrated LITHOPROBE seismic reflection sections. See legend and geological map of Fig. 1 for location of seismic lines and explanation of most abbreviations. Line 12 was recorded with regional parameters and line 12A with high-resolution parameters (see Fig. 2 legend for details). The panel for line 14 includes information from regional seismic section 14 and two high-resolution sections: 14A recorded across the PDF (this 6-km portion of data was not migrated) and 14B along the south half of line 14. M, probable Moho; BR, Blake River group; PS, Pontiac subprovince; other abbreviations are referred to in text.

the gently dipping reflection zones can be tied to surface features. But age inversions of metavolcanic rock assemblages a little to the south of line 12A suggest imbrication of some crustal units²¹. In conjunction with the new age data, the character of the north-dipping layered reflections beneath line 12A thus supports a model of south-verging thrust faulting.

During the past decade, the principal structures in the southern AGB have been interpreted in terms of (1) north-south directed compression^{8,10}, (2) sinistral transpression^{20,22}, (3) extension^{5,10,22}, or (4) crustal-scale subsidence⁹, with the LCF and PDF being explained variously as normal (growth) and/or thrust and/or strike-slip fault zones. Our interpretation of the seismic reflection data favours interpretations 1 and 2. We have found no evidence for structures of the type recently imaged across the North American mid-continent rift system (for example, fanning of layered strata and growth faults)²³, which would be expected if the AGB volcanic rocks had been deposited in subsiding grabens or half grabens with only minor subsequent orogeny^{5,9}. Although magma composition suggests that extension played a key role in early AGB magmatism and tectonism²², it is evident from the seismic reflection sections and the new geochronological data²¹ that extensional structures were

largely overprinted by the effects of later compression and a significant, but undetermined, component of strike-slip movement.

Some ambiguity in the relative age of thrust and strike-slip faulting remains. Orogeny either involved early folding and thrust faulting with structures dissected by later strike-slip faults or there were concurrent thrust and strike-slip movements in a transpressional regime. Post-tectonic plutons emplaced in the southern AGB and adjacent Pontiac subprovince are only 10–15 Myr younger than the principal metavolcanic rocks in this region^{21,24}. Such a short period from voluminous magmatism and tectonism involving thrust and strike-slip faulting to stabilization of the Archaean craton is compatible with an accretionary regime in which neighbouring volcanic and sedimentary terranes were overthrust and juxtaposed laterally in a zone of oblique convergence. □

Received 19 December 1989; accepted 15 February 1990.

1. McLennan, S. M. *Pure appl. Geophys.* **128**, 683–724 (1988).
2. Condie, K. C. *Archaean Greenstone Belts* (Elsevier, Amsterdam, 1981).
3. Windley, B. F. *The Evolving Continents* 2nd edn (Wiley, Chichester, 1984).
4. Kroner, A. *Phys. Chem. Earth* **15**, 69–106 (1985).
5. Gibson, I. L., Roberts, R. G. & Gibbs, A. *Earth planet. Sci. Lett.* **79**, 159–167 (1986).

6. Condie, K. C. *Precamb. Res.* **32**, 261–278 (1986).
7. Hoffman, P. F. in *Decade of North American Geology* Vol. A (eds Bally, A. W. & Palmer, A. R.) 447–512 (Geological Society of America, Boulder, Colorado, 1989).
8. Dimroth, E. et al. *Can. J. Earth Sci.* **20**, 1374–1388 (1983).
9. Jensen, L. S. *Spec. Pap. geol. Ass. Can.* **28**, 65–87 (1985).
10. Hodgson, C. J. *Trans. Inst. Miner. Metall.* **95**, B183–194 (1986).
11. Percival, J. A. & Card, K. D. *Geology* **11**, 323–326 (1983).
12. Percival, J. A. et al. *Nature* **342**, 416–419 (1989).
13. Barley, M. E. et al. *Geology* **17**, 826–829 (1989).
14. Gibbs, A. K. et al. *Geol. Soc. Am. Bull.* **95**, 280–294 (1984).
15. Gibbs, A. K. et al. *Am. geophys. Un. Geodyn. Ser.* **14**, 95–106 (1986).
16. Boland, A. et al. *Nature* **335**, 711–713 (1988).
17. Wilson, M. E. *Mem. geol. Surv. Can.* **315** (1962).
18. Fountain, D. M., Salisbury, M. H. & Percival, J. A. *J. geophys. Res.* (in the press).
19. Parker, C. L. thesis, McGill Univ. (1984).
20. Hubert, C., Trudel, P. & Gelin, L. *Can. J. Earth Sci.* **21**, 1024–1032 (1984).
21. Corfu, F. et al. *Can. J. Earth Sci.* **26**, 1747–1763.
22. Ludden, J. & Hubert, C. *Geology* **14**, 707–711 (1986).
23. Cannon, W. C. et al. *Tectonics* **8**, 305–322 (1989).
24. Mortensen, J. K., Theriault, R. J. & Card, K. C. *Geol. Ass. Can. A Meet. Prog. Abstr.* **13**, A87 (1988).

ACKNOWLEDGEMENTS. We thank Drs K. D. Card and J. A. Percival for reviews of an earlier draft of this manuscript, P. Hurlay, G. Gauthier and numerous others involved in the acquisition of the seismic data, and Veritas Geophysical Ltd and Veritas Seismic Ltd for collecting and processing our data. The 1987–88 LITHOPROBE survey was funded by the Natural Science and Engineering Research Council of Canada, Geological Survey of Canada, Ontario Geological Survey, Ministère de l'Energie et des Ressources du Québec, and Minnova Inc.

Female sticklebacks use male coloration in mate choice and hence avoid parasitized males

Manfred Milinski & Theo C. M. Bakker

University of Bern, Zoologisches Institut, Abteilung Verhaltensökologie, Wohlenstrasse 50a, CH-3032 Hinterkappelen, Switzerland

AN important problem in evolutionary biology since the time of Darwin has been to understand why females preferentially mate with males handicapped by secondary sexual ornaments^{1–3}. One hypothesis of sexual selection theory is that these ornaments reliably reveal the male's condition^{4–6}, which can be affected for example by parasites^{4,7–13}. Here we show that in the three-spined stickleback (*Gasterosteus aculeatus*) the intensity of male red breeding coloration positively correlates with physical condition. Gravid females base their active mate choice on the intensity of the male's red coloration. Choice experiments under green light prevent the use of red colour cues by females, and males that were previously preferred are now chosen no more than randomly, although the courtship behaviour of the males remains unchanged. Parasitization causes a deterioration in the males' condition and a decrease in the intensity of their red coloration. Tests under both lighting conditions reveal that the females recognize the formerly parasitized males by the lower intensity of their breeding coloration. Female sticklebacks possibly select a male with a good capacity for paternal care¹⁴ but if there is additive genetic variation for parasite resistance, then they might also select for resistance genes, as proposed by Hamilton and Zuk⁴.

At the start of the breeding season, male three-spined sticklebacks develop a bright red coloration due to carotenoids¹⁵, and it has been shown that females prefer artificially coloured males over colourless males¹⁶. In another fish species, the guppy (*Poecilia reticulata*), female choice is based not only on the expression of these pigments^{17–21}, which may be indicative of fitness¹⁷, but also on courtship behaviour^{22,23}.

Twenty-four male three-spined sticklebacks with developed breeding coloration were placed individually into tanks (17.8 cm × 34.5 cm, with a water level of 16.5 cm and at a temperature of about 18 °C) separated by grey opaque partitions. Each pair of tanks was illuminated for 16 hours per day by a 60 W reflector bulb (Osram Concentra PAR-EC). Each male was stimulated with a ripe female enclosed in a plexiglass cell (11 cm × 7.5 cm, water level 16 cm) placed close to the front wall

of its tank for five minutes daily to accelerate its nestbuilding behaviour²⁴. After six days, all the males had a complete nest built in a Petri dish provided close to the backwall and were courting vigorously.

Two students scored the intensity of the red breeding coloration on a 10-point scale (1, duller male; 10, brightest male) for each male when it courted a female. There was general agreement between the students ($r = 0.71$). Males designated 1 and 2, 3 and 4, and so on, according to increasing colour rank, were defined as pairs for presentation to ripe females. To avoid the right male always being brighter, positions were randomized within pairs. In a separate tank positioned centrally in front of each pair of neighbouring tanks, the cell containing a single gravid female was placed; her choosing process between the two males was video-recorded for a 5-min period after 1 min of acclimatization. Females were previously selected for their readiness to spawn, that is, to adopt and maintain the head-up courtship posture while pointing towards one of the two males. On each day before we gave a female the opportunity to choose between males, we estimated the difference of red colour between the members of each pair on a 5-point scale (0, no difference; 1, slight but distinct difference; 2, pronounced difference, and intermediates). Each of four females chose between each of the 12 pairs of males (female 1 made 12 choices on day 1, female 2 made 12 choices on day 2, and so on). The

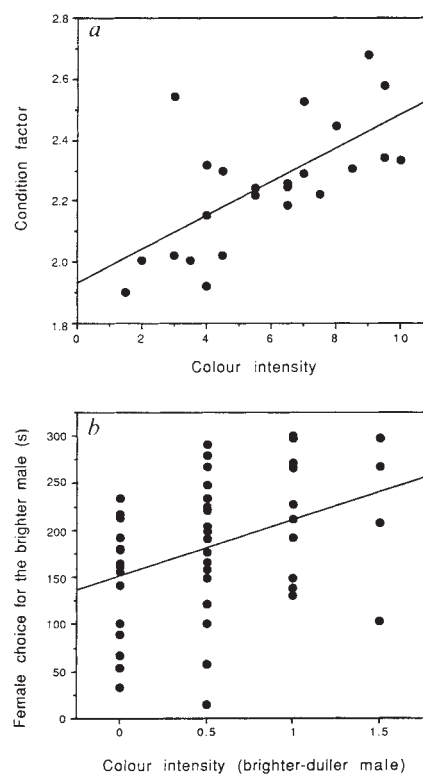


FIG. 1 a, Correlation between the intensity of red breeding coloration (average score of 2 students) and the condition factor ($100 \times \text{weight (g)} / \text{length (cm)}^{2.76}$) of 24 reproductive male sticklebacks ($y = 1.93 + 0.056x$, $r^2 = 0.44$, $F = 17.27$, d.f. = 1,22, $P < 0.0004$, 1-tailed). The condition factor, used as a standard practice in fisheries ecology, is regarded as a good indicator of the general well-being of teleost fishes³⁷. It assumes that heavier fish of a given length are in better condition. b, Correlation between the difference in colour intensity scored daily of 12 pairs (brighter–duller male) and active female choice for the brighter male (measured in seconds, see text) by 4 different females, one on each day. The pooled regression is significant ($y = 151.0 + 59.1x$, $r^2 = 0.14$, $F = 7.62$, d.f. = 1,46, $P < 0.004$, 1-tailed). Pooling of females is allowed for²⁸ and recommended³⁸, because the slopes of regression lines of the 4 females were not significantly different ($F = 0.60$, d.f. = 3,40, $P = 0.62$, 2-tailed) and the subsequent analysis of covariance revealed no significant differences either ($F = 2.13$, d.f. = 3,43, $P = 0.11$, 2-tailed).

intensity of the red breeding coloration correlated significantly positively with the males' condition factor (Fig. 1a). Therefore, females could judge the well-being of a male from the intensity of its red coloration. Although the experiment was designed so that the males of each pair were very similar in coloration, there was a significant preference by the females for the brighter male in each pair, this preference being intensified as the difference in coloration increased (Fig. 1b).

To investigate whether this preference is ultimately based on coloration or on some related character such as courtship behaviour, we repeated the same experiment using 15 new pairs of males; females could choose between males under white light and under such light conditions that they were almost unable to assess differences in the intensity of red coloration (males and females had not been used in the previous experiment). To achieve this, males and females were kept under white and green light (Osram Concentra PAR-EC Belcolor, 80 W) on alternate days. After the 30 males had built nests in their individual tanks, we ordered the tanks according to the intensity of the inhabitant's red coloration; six students then estimated the difference in red colour between the members of each pair on the 5-point scale (median value of all possible correlations among the students, $r=0.77$). The difference in intensity of coloration (brighter-duller male) correlated positively with the difference in condition (Fig. 2a).

Under white light, each of 13 different ripe females was allowed to choose between males from as many different pairs as she was reliably willing to (3.5 pairs per female on average), as determined by her head-up posture during the whole trial. Hence, the first female was allowed to choose between the males of the duller pair, thereafter between the males of the brighter neighbouring pair and so on, until she had to be replaced by the second female, and so on. Under green light, 11 other females were used (4.0 pairs per female). Thus, each pair of males was confronted with three different females in either experiment. To establish whether the courtship behaviour of the males varies under the different light conditions, we confronted each male alone with the cell containing a ripe female under both green and white light for a 4-min period. The number of zigzags performed in the male's courtship display²⁵, regarded as a reliable measure of male courtship intensity^{26,27}, was counted during the last three minutes. The males' courtship intensity did not differ significantly in the two lighting situations (mean number of zigzags under white light, 46.5, s.d. = 28.9; under green light, 42.1, s.d. = 25.7; $P>0.10$, Wilcoxon matched-pairs signed-ranks test, 2-tailed) and differences between pair members did not change significantly (zigzags for the brighter male of a pair, 61.5%, s.d. = 28.6, under white light; and 53.0%, s.d. = 29.2, under green light; $P>0.10$, Wilcoxon matched-pairs signed-ranks test, 2-tailed). Also the females' willingness to react to male courtship as depicted by duration of head-up posture was not influenced by the type of light (3.5 pairs per female and 4.0 pairs per female, respectively; see above). A female that was allowed to enter the male's territory under green light went through the normal spawning sequence immediately.

Under white light, females preferred the redder male again (Fig. 2b), whereas under green light the trend in favour of the brighter males was not significant (Fig. 2c); the slope of the regression under white light is significantly greater than that under green light ($F=2.99$, d.f. = 1,26, $P=0.05$, 1-tailed). This indicates that the females based their choice primarily on differences in male red coloration. This makes sense functionally, because colour intensity correlates significantly with condition (see Fig. 2a; partial correlation²⁸ when courtship intensity was kept constant, $r=0.54$, $P<0.03$, 1-tailed), whereas courtship intensity does not correlate significantly with condition ($r=0.00$, $P>0.10$, 1-tailed; partial correlation when colour intensity kept constant, $r=-0.23$, $P>0.10$, 1-tailed). The conclusion that female choice is based mainly on colour cues is confirmed by partial correlations under white light. The

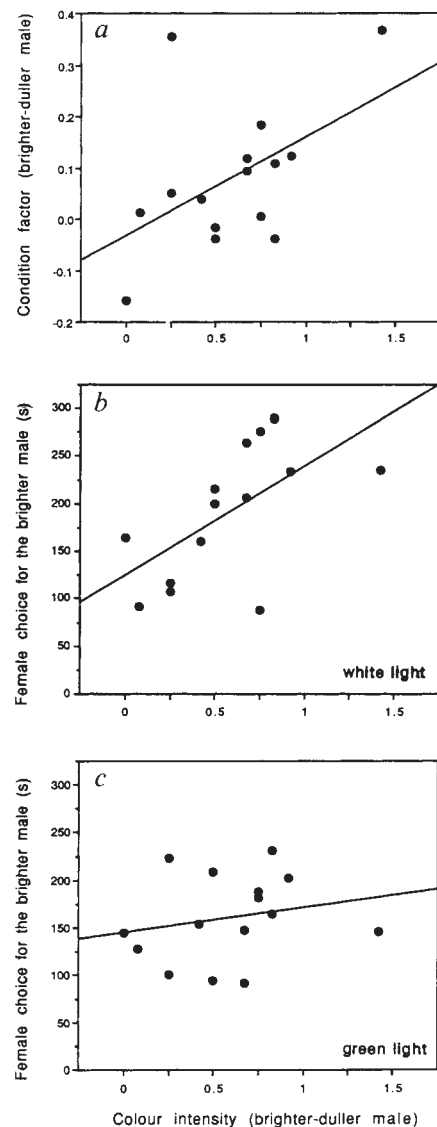


FIG. 2 Difference in the intensity of red breeding coloration of 15 pairs of reproductive male sticklebacks (brighter-duller male) (average score of 6 students) in relation to: a, the difference in condition factor ($100 \times \text{weight (g)} / \text{length (cm)}^{2.89}$) ($y = -0.034 + 0.192x$, $r^2=0.25$, $F=4.25$, d.f. = 1,13, $P<0.03$, 1-tailed); b, average active female choice of the brighter male by 3 different females per pair (13 females in total) under white light ($y = 125.09 + 120.2x$, $r^2=0.38$, $F=7.97$, d.f. = 1,13, $P<0.01$, 1-tailed); c, average active female choice of the brighter male by 3 different females per pair (11 females in total) under green light ($y = 144.80 + 26.17x$, $r^2=0.04$, $F=0.60$, d.f. = 1,13, $P=0.75$, 2-tailed). The 15 average choices under either light condition are treated as independent sample units because the experiment shown in Fig. 1b did not reveal significant differences among females choosing between 12 pairs of males. To test whether the intensity of the green light (330 lx at the bottom of the tank) was sufficient to prevent females from dropping below the point at which the Purkinje shift occurs, when their ability to distinguish different intensities of red, for example, would be impeded, we submitted three females to a discrimination test under white light at a lowered intensity of 160 lx. Two red colour cards (1.7×2.1 cm) of the same hue and grey value, but of different colour intensities (Munsell System no. 7.5 R, 5/14 and no. 7.5 R, 5/2) were presented simultaneously to the fish and positions were alternated randomly. Approaching the less-red card was rewarded with food. In 50 training sessions, two fish achieved 80% and one 100% correct responses (approaching the less-red card within a distance of 5 cm before a reward) in the last 10 trials (as $\chi^2=2.31$, d.f. = 2, $P=0.32$, 2-tailed, reveals homogeneity among the 3 females, the 30 choices were pooled, 26:4 was significantly different from 1:1, $\chi^2=16.13$, d.f. = 1, $P<0.001$, 1-tailed). Thus, even at half the intensity of the green light, females could discriminate between different intensities of red colour when in white light.

correlation between choice and intensity of red coloration when intensity of courtship is kept constant is $r = 0.54$ ($P < 0.03$, 1-tailed), whereas the almost significant correlation between choice and intensity of courtship ($r = 0.42$, $P = 0.06$, 1-tailed) is far from significant when intensity of red coloration is kept constant ($r = 0.25$, $P > 0.10$, 1-tailed). Also, in guppies no correlation between orange breeding coloration and display rate has been found²⁰, but contrary to our results, female guppies respond not only to the colour of the males' orange spots, but also to their contrast against background skin under filtered light²⁹.

To investigate whether parasites influence both the males' red breeding coloration and the result of active female choice, we infested the brighter male of each pair with the ciliate *Ichthyophthirius multifiliis*, a serious and widespread fish disease known as 'white-spot'³⁰. On five successive days, 50 ml of water contaminated with tomites (the infective stage) were taken from the tank of a formerly heavily infected stickleback and poured twice daily into the tank of each fish. Uncontaminated aliquots of water were added to the tanks of control fish. After several days, infected fish developed from a few to about 50 visible white cysts, which dropped off the fish after a few days. Two fish died after infection. To prevent reinfection, all tanks (including those of control fish) were treated with *Faunamor*. Surviving fish continued to court stimulus females. Four of the six students

who had scored the males' red coloration were asked to repeat their estimation of pair differences; they were not informed about the earlier parasitization of half of the males. Once more, ripe females which had not been used in previous experiments were allowed to choose between males of the remaining 13 pairs both under white (3 different females) and green light (4 other females) so that each pair of males was presented with one female in each experiment.

Parasitization caused a significant decrease in the intensity of the males' red coloration (Fig. 3a) and in their condition factor (Fig. 3b). Females significantly reduced their earlier preference for the males that were formerly brighter under white light (Fig. 3c), but under green light the males' parasitization had no significant effect on female choice (Fig. 3d). This implies that the females detected the prior parasitization of the males by their decreased intensity of breeding coloration, which is a necessary condition for coloration to be judged as a revealing handicap^{4,6}.

Under natural conditions, brighter males might obtain better territories by dominance interactions, a factor that was excluded in this study. Intersexual selection on male stickleback red breeding coloration seems, however, to be more important than intrasexual selection, because male sticklebacks develop more erythrophores³¹ and 'flush' their red colours more strongly³² after presentation with a ripe female than after exposure to a rival male. Furthermore, the overall intensity and intermale variation of red coloration is greatest during the courtship stage of the breeding cycle³³. By contrast to the males, the females' visual sensitivity for red coloration periodically increases at the beginning of the reproductive season and reaches a higher level than that of males³⁴.

The females probably did not make use of the male's courtship intensity for their decision-making because courtship intensity is a poor predictor of condition. Perhaps even a sick or convalescent male can muster energy for the display when need arises, but it is harder for him to bluff the long-term drain on his resources revealed by his lack of colour. Nevertheless, the zigzag display may help the females to recognize a reproductive male stickleback.

In all, the intensity of the red breeding coloration seems to be a revealing handicap^{4,6} for a male's condition because it correlates significantly positively with the condition of our wild-caught males and decreases when the males' condition is experimentally reduced by parasitization. Therefore, any agent (including parasites) influencing a male stickleback's condition probably affects the intensity of its breeding coloration and consequently the female's choice. Why do female sticklebacks prefer males of superior condition? As the male cares for the eggs and the fry for about 10 days after spawning²⁶, she might prefer a strong male with a high probability of survival for this period¹⁴. Even if she prefers a strong male for paternal care, she cannot avoid simultaneously selecting for genes favourable for parasite resistance if there is additive genetic variation for parasite resistance in the population. Although present evidence is ambiguous with respect to *I. multifiliis*³⁵, there are indications of such a variation in fish³⁶. Therefore the Hamilton-Zuk process⁴ may be an auxiliary factor in species with paternal care. □

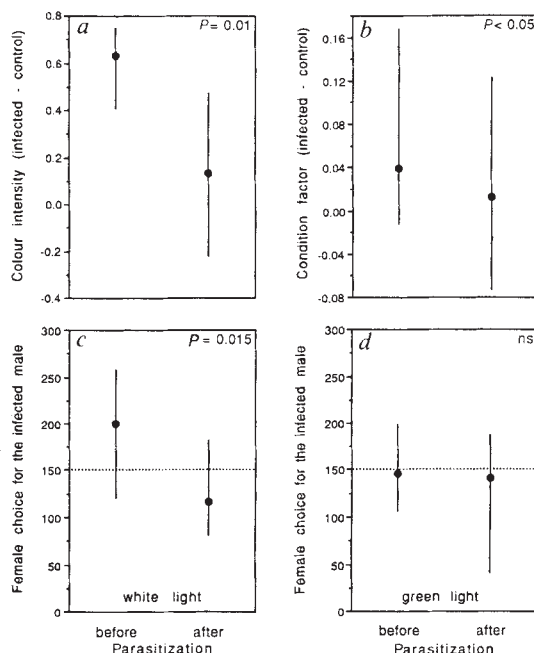


FIG. 3 Median difference of 13 pairs of reproductive male sticklebacks (infected - control) before parasitization of the brighter male with *I. multifiliis* and after recovery from the parasite a, in the intensity of red breeding coloration (average score of 4 students) before (median correlation between students, $r = 0.60$) and after infection (median $r = 0.59$) (the same students had difficulties in finding differences in intensity of the red coloration between the males after the more brightly coloured male of each pair had been parasitized. As the students' ability to detect differences should have improved with experience, the students' second scoring was conservative); and b, in the condition factor ($100 \times \text{weight (g)} / \text{length (cm)}^{2.89}$). Median active female choice of the infected male of each pair before parasitization with *I. multifiliis* and after recovery from the parasite c, by 3 different females (1 female per pair) under white light, and d, by 4 different females (1 female per pair) under green light. Before parasitization females chose the brighter males significantly longer under white light than under green light ($P < 0.01$, Wilcoxon matched-pairs signed-ranks test, 1-tailed). After parasitization the respective difference was not significant ($P > 0.10$, Wilcoxon matched-pairs signed-ranks test, 2-tailed). This is expected because both the colour difference (Fig. 3a) and the condition difference (Fig. 3b) between the males had almost disappeared after parasitization of the brighter male. Bars give quartiles, dotted line indicates no preference, probabilities after Wilcoxon matched-pairs signed-ranks tests, 1-tailed, N.S. $P > 0.10$.

Received 2 October; accepted 18 December 1989.

- Bradbury, J. W. & Andersson, M. B. (eds) *Sexual Selection: Testing the Alternatives* (Wiley, New York, 1987).
- Kirkpatrick, M. *Ann. Rev. Ecol. Syst.* **18**, 43-70 (1987).
- Maynard-Smith, J. *J. theor. Biol.* **115**, 1-8 (1985).
- Hamilton, W. D. & Zuk, M. *Science* **218**, 384-387 (1982).
- Zahavi, A. *J. theor. Biol.* **53**, 205-214 (1975).
- Andersson, M. *Biol. J. Linn. Soc.* **17**, 375-393 (1982).
- Read, A. F. *Nature* **328**, 68-70 (1987).
- Ward, P. I. *Anim. Behav.* **36**, 1210-1215 (1988).
- Ward, P. I. *Oikos* **55**, 428-429 (1989).
- Read, A. F. & Harvey, P. H. *Nature* **339**, 618-620 (1989).
- Pomiankowski, A. *Nature* **338**, 115-116 (1989).
- Endler, J. A. & Lyles, A. M. *Trends Ecol. Evol.* **4**, 246-248 (1989).

13. Read, A. *Trends Ecol. Evol.* **3**, 97–101 (1988).
14. Heywood, J. S. *Evolution* **43**, 1387–1397 (1989).
15. Brush, A. H. & Reisman, H. M. *Comp. Biochem. Physiol.* **14**, 121–125 (1965).
16. Semler, D. E. *J. Zool.* **165**, 291–302 (1971).
17. Endler, J. A. *Evolution* **34**, 76–91 (1980).
18. Endler, J. A. *Env. Biol. Fish.* **9**, 173–190 (1983).
19. Kodric-Brown, A. *Behav. Ecol. Sociobiol.* **17**, 199–205 (1985).
20. Houde, A. E. *Evolution* **41**, 1–10 (1987).
21. Houde, A. E. *Anim. Behav.* **36**, 510–516 (1988).
22. Kennedy, C. E. J., Endler, J. A., Poynton, S. L. & McMinn, H. *Behav. Ecol. Sociobiol.* **21**, 291–295 (1987).
23. Bischoff, R. J., Gould, J. L. & Rubenstein, D. I. *Behav. Ecol. Sociobiol.* **17**, 253–255 (1985).
24. Wootton, R. J. *The Biology of the Sticklebacks* (Academic, London, 1976).
25. ter Pelkewijk, J. J. & Tinbergen, N. *Z. Tierpsychol.* **1**, 193–200 (1937).
26. van Iersel, J. J. A. *Behaviour Suppl.* **3**, 1–159 (1953).
27. Sevenster, P. *Behaviour Suppl.* **9**, 1–170 (1961).
28. Sokal, R. R. & Rohlf, F. J. *Biometry* 2nd edn (Freeman, New York, 1981).
29. Long, K. D. & Houde, A. E. *Ethology* **82**, 316–324 (1989).
30. Smyth, J. D. *Introduction to Animal Parasitology* (Hodder and Stoughton, London, 1985).
31. Reisman, H. M. *Copeia* **1968**, 816–826 (1968).
32. Bakker, T. C. M. *Behaviour* **98**, 1–144 (1986).
33. McLennan, D. A. & McPhail, J. D. *Can. J. Zool.* **67**, 1767–1777 (1989).
34. Cronly-Dillon, J. & Sharma, S. C. *J. exp. Biol.* **49**, 679–687 (1968).
35. McCallum, H. I. *Parasitology* **85**, 475–488 (1982).
36. Price, D. J. *J. Fish Biol.* **26**, 509–519 (1985).
37. Bolger, T. & Connolly, P. L. *J. Fish Biol.* **34**, 171–182 (1989).
38. Miller, R. G. Jr *Beyond ANOVA, Basics of Applied Statistics* (Wiley, New York, 1986).

ACKNOWLEDGEMENTS. We thank J. A. Endler, W. D. Hamilton, P. H. Harvey, L. Partridge, M. Petrie and M. Zuk for their advice on the manuscript, H. Riedwyl for most of the statistical work, O. Lassière for improving our English and eight students for scoring the sticklebacks' breeding coloration.

Effects on lower trophic levels of massive fish mortality

Michael J. Vanni*, Chris Luecke*, James F. Kitchell, Yvonne Allen, Jo Temte & John J. Magnuson

Center for Limnology, University of Wisconsin, Madison, Wisconsin 53706, USA

PREDATION can be a potent force structuring ecological communities and affecting several trophic levels^{1–7}. The cascading trophic interactions hypothesis predicts that lacustrine predators such as fish can have a strong effect on herbivorous zooplankton, which in turn can regulate phytoplankton^{8,9}. Ascertaining the scale, scope and generality of this hypothesis is important for both development of ecological theory and aquatic ecosystem management¹⁰. Although small-scale tests of parts of this cascade are common^{11–15}, whole-lake assessments are not, particularly in large, natural lakes⁹; also the influence of several trophic levels and nutrients, a requisite for unambiguous interpretation of predator effects, has not been considered. Here we present data from a natural experiment in Lake Mendota, Wisconsin, USA, which support the cascading trophic interactions hypothesis. Massive mortality of fish greatly reduced predation on zooplankton, resulting in an increase in the abundance of large *Daphnia*, and a dramatic decrease in phytoplankton biomass. Physical factors and concentrations of limiting nutrients were unchanged before and after fish mortality, indicating that these factors probably did not cause the observed decrease in phytoplankton. Our results demonstrate strong food web influences on phytoplankton, and support the idea that food web interactions can be managed to reduce phytoplankton abundance.

In Lake Mendota, Wisconsin, an unusually warm summer (1987) resulted in massive mortality of planktivorous fish. We have compared the relative abundance of fish, zooplankton, phytoplankton and dissolved nutrients before and after fish mortality, to assess the effects fish can have on freshwater planktonic communities.

Seasonal dynamics of the Lake Mendota food web were relatively consistent from 1978–1986 (refs 16, 17). Dominant planktivorous fish were cisco, *Coregonus artedii*, and yellow

TABLE 1 Nutrient limitation indicators of Lake Mendota phytoplankton

Experiment	Biomass response (treatment/control)	Nutrient uptake rate*	
		Phosphorus	Nitrogen
Before cisco die-off			
24–28 July 1986	4.3	1.3	14.6
21–25 August 1986	4.8	1.5	16.6
4–8 August 1987	3.2	ND	ND
After cisco die-off			
7–11 July 1988	1.0	0.8	5.0
5–8 August 1988	0.9	1.4	5.3

The table demonstrates that Lake Mendota phytoplankton were more nutrient-limited before rather than after the cisco die-off. Nutrient limitation was estimated in two ways in each experiment: (1) increase in phytoplankton biomass when nutrients were added, relative to controls with no nutrient additions; and (2) rates of nutrient uptake by phytoplankton. In both cases, higher values indicate more severe nutrient limitation. Lake phytoplankton were incubated in clear plastic enclosures (19 litres) at 50% surface light intensity. Nutrients (80 µM nitrogen as NH₄NO₃ and 3.5 µM phosphorus as KH₂PO₄) were added to some enclosures and others served as controls. Enclosures were sampled 3–4 days later for phytoplankton biomass (chlorophyll *a* concentration) and nutrient concentrations¹⁷. To calculate uptake rates, it was assumed that all N and P added to enclosures and not remaining in solution was taken up by phytoplankton. Biomass response refers to the final concentration of chlorophyll *a* in treatment enclosures (nitrogen and phosphorus added) divided by chlorophyll *a* in control enclosures (no nutrients added). ND, indicates that nutrient uptake rates were not determined.

* Nutrient uptake rate measured as µM per µg of chlorophyll per day to standardize for phytoplankton biomass.

perch, *Perca flavescens*¹⁸. Dominant zooplankton were cyclopoid copepods in early spring, *Daphnia galeata mendotae* in late spring, and cyclopoid and calanoid copepods in summer^{16,19}. Diet analyses and bioenergetics modelling showed that cisco was the dominant zooplanktivore in spring, accounting for >50% of predatory mortality of *Daphnia*. Both perch and cisco were important zooplanktivores in summer, and both selectively prey on *Daphnia* over other zooplankton¹⁸. Phytoplankton seasonal succession during the years 1978–1986 was typical of eutrophic lakes^{16,17}. After ice-out, a bloom of diatoms and flagellates developed, followed by a clear water period²⁰ (low phytoplankton biomass and high water transparency). *Daphnia* declined in late spring, and remained rare through the summer, during which time phytoplankton biomass was high and dominated by blue-green algae¹⁶. The spring clear water period was always associated with the annual maximum abundance of *Daphnia galeata mendotae*¹⁶. Experiments showed that *Daphnia* have higher grazing rates than other Lake Mendota zooplankton and that *Daphnia* grazing causes the clear water period¹⁷.

In 1987 we initiated a study of the Lake Mendota food web, with the aim of understanding the factors controlling the seasonal dynamics of plankton, and applying them in the effective management of algal biomass. We monitored all trophic levels of the planktonic food web and potential limiting nutrients, during 1987 and 1988. In late August and September 1987, a massive mortality of cisco occurred¹⁸. This die-off, combined with modelling and smaller-scale experimentation, presented an opportunity to observe and interpret the effects of predator reduction on lower trophic levels. Almost the entire cisco population died during this time, as evidenced by gill-net catches (see legend to Fig. 1) and acoustic estimates of cisco abundance, which showed that cisco biomass declined from 239 kg per hectare in 1987 to 13 kg per hectare in 1988 (P. Jacobson and J.J.M., unpublished data). Cisco mortality was probably due to unusually warm epilimnetic temperatures and depletion of hypolimnetic oxygen.

Effects of the cisco die-off on *Daphnia* were dramatic (Fig. 1). The large-bodied *Daphnia pulicaria*, extremely rare in the lake for the previous decade, became abundant and replaced the smaller *D. galeata mendotae*. The increase in *Daphnia* biomass and body size resulted in a 2–3-fold increase in grazing rates on phytoplankton (average from April–September), based on the relationship between body size and individual grazing rate²¹. The *Daphnia* grazing rate for April 1988 was ten times that of the April 1987 *Daphnia* grazing rate.

* Present addresses: Department of Zoology, Miami University, Oxford, Ohio 45056, USA (M.J.V.), and Department of Fisheries and Wildlife, Utah State University, Logan, Utah 84322, USA (C.L.).

Response of the phytoplankton community to the cisco die-off was equally striking (Fig. 1). In 1987, the spring phytoplankton bloom developed shortly after ice-out and persisted through to mid-May. *Daphnia galeata mendotae* then increased in abundance, causing the clear water period (Fig. 1). Following the decline of *Daphnia galeata* in 1987, the summer phytoplankton bloom developed, consisting of blue-green algae and the dinoflagellate *Ceratium*. Spring phytoplankton biovolume was much lower in 1988, corresponding with a greatly increased *Daphnia pulicaria* abundance (Fig. 1). The spring clear water period (phytoplankton biovolume $<1.5 \times 10^6 \mu\text{m}^3 \text{ml}^{-1}$) lasted nearly 6 weeks in 1988, compared with 1–2 weeks in 1987 and previous years since 1978 (ref. 16). As in 1987, phytoplankton increased rapidly in late May–early June 1988 (Fig. 1); however, in 1988 phytoplankton quickly declined, and remained relatively low through the summer (Fig. 1).

Differences in phytoplankton dynamics between 1987 and 1988 were probably not caused by differences in epilimnetic concentrations of potential limiting nutrients (Fig. 2). From ice-out through May, concentrations of soluble reactive phosphorus were actually higher in 1988 than 1987, and concentrations of nitrate/nitrite were higher from ice-out through July in 1988 than they were in 1987 (Fig. 2). Concentrations of ammonia were more variable, but closely correlated with *Daphnia* biomass in both years; these results and other data¹⁷ show that *Daphnia* excrete considerable quantities of ammonia. Thus *Daphnia pulicaria* increases dissolved nutrient concentrations in two ways: (1) by reducing phytoplankton biomass, and therefore total uptake of dissolved nutrients and (2) by direct excretion.

Evaluation of difference in dissolved nutrients in the summers of 1987 and 1988 is difficult because their concentrations were at or near analytical detection limits (Fig. 2). But we assessed nutrient limitation by measuring the response of phytoplankton to experimental addition of nutrients. Nutrient additions stimulated phytoplankton growth and nutrient uptake rates more before the cisco die-off than after the die-off (Table 1). Both responses suggest that the demand for nutrients was less in summer 1988 than it was in summer 1987.

Epilimnetic temperatures and hypolimnetic nutrient con-

TABLE 2 Monthly epilimnetic temperatures and dissolved hypolimnetic nutrient concentrations in Lake Mendota

	Epilimnion temperature (°C)		
	June	July	August
1987	21.7	23.3	23.3
1988	19.5	23.6	24.2
	Dissolved inorganic N (μM)		
	Early July	Early August	Late August
1987	41.1	47.9	39.6
1988	60.0	53.9	54.3
	Soluble reactive P (μM)		
	Early July	Early August	Late August
1987	4.1	4.9	3.5
1988	4.4	5.8	3.3

The table shows that epilimnetic temperatures and dissolved hypolimnetic nitrogen and phosphorus concentrations were similar in 1987 and 1988. Each epilimnetic temperature value represents the mean temperature of several depths from 0–10 m, averaged over several dates for each month. Each nutrient concentration represents the mean concentration on a single date at 12 and 14 m. These depths are just below the epilimnion and constitute strata most likely to be entrained into the epilimnion. Dissolved inorganic N is the sum of ammonium and nitrate/nitrite concentrations. Nutrient concentrations were assayed using standard spectrophotometric methods²³.

centrations were fairly similar in 1987 and 1988 (Table 2), although hypolimnetic nutrient concentrations tended to be slightly higher in 1988. These and the nutrient limitation data indicate that temperature differences, and their potential effects on entrainment of nutrients from the hypolimnion to the epilimnion, were probably not responsible for the 1988 decline in phytoplankton.

Events following the 1987 cisco die-off in Lake Mendota support the cascading trophic interactions hypothesis, which predicts large *Daphnia* to be a key species⁹. Cisco die-offs have

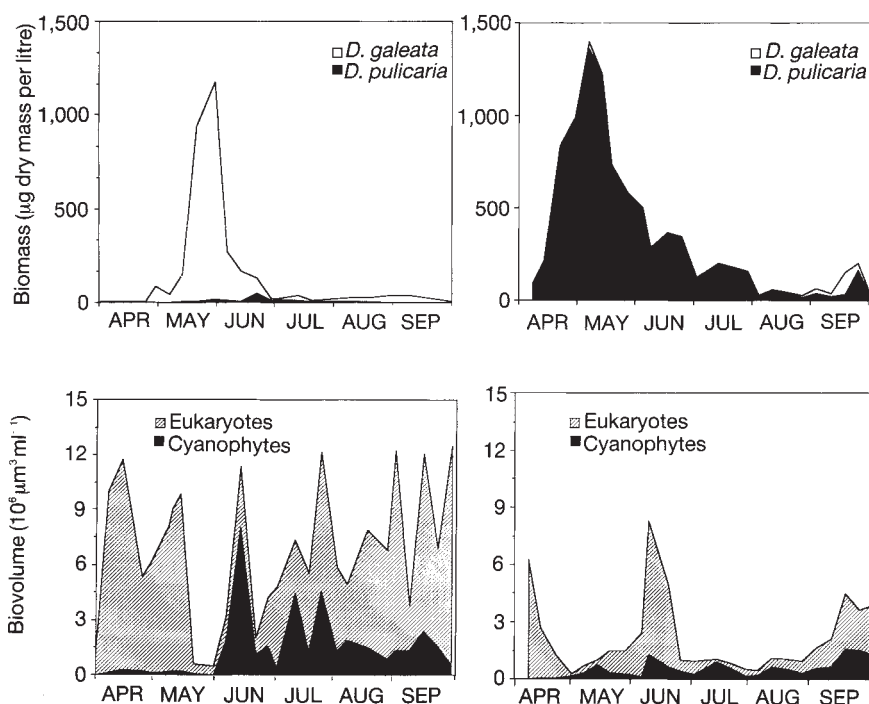


FIG. 1 Abundance of *Daphnia* and phytoplankton in Lake Mendota in 1987 (left panels) and 1988 (right panels). Planktivorous fish (cisco and yellow perch) were sampled monthly in the pelagic area with vertical (0–20 m) gill nets. One net of each of the following sizes (stretched mesh) was used: 19, 25, 32, 38, 51, 65, 89, 125 mm. The biomass (g wet mass per net per day) of cisco declined from 860 in 1987 to less than 20 in 1988; perch biomass increased from 250 in 1987 to 460 in 1988. Zooplankton were sampled at four depth strata (0–5, 5–10, 10–15 and 15–20 m) with a Clarke–Bumpus sampler; water column density was calculated as the average of the four strata. *Daphnia* biomass was obtained using length–weight regressions²⁴. Phytoplankton were sampled with a tube that collected a 0–10 m integrated sample, corresponding to the depth of the mixed layer; cells were counted with an inverted microscope¹⁷.

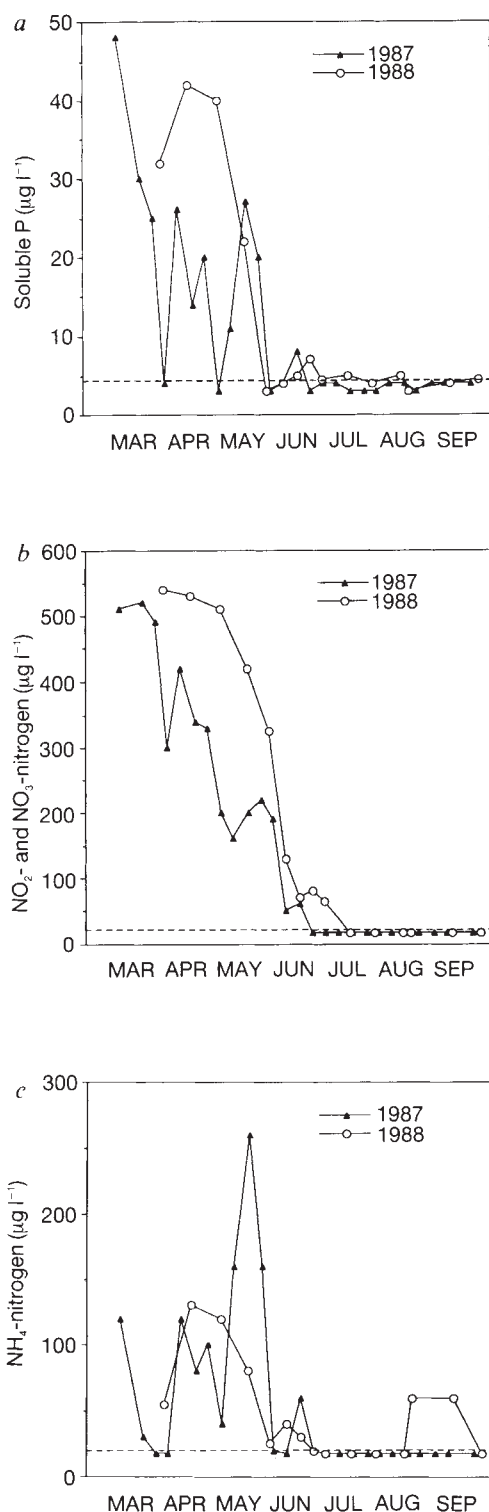


FIG. 2 Concentrations of dissolved phosphorus and nitrogen in Lake Mendota in 1987 and 1988: *a*, Soluble reactive phosphorus; *b*, nitrite and nitrate; *c*, ammonia. Dashed lines indicate analytical detection limits. Nutrients were sampled from the epilimnion (0–10 m integrated or 4 m); analyses were conducted by the Environmental Sciences Section of the Wisconsin State Laboratory of Hygiene, using standard analytical methods²³.

occurred before in Lake Mendota²², and sparse cisco populations in the 1970s are associated with the dominance of *Daphnia pulicaria* and relatively long spring clear-water periods¹⁶. Thus, events in Lake Mendota provide a dramatic large-scale demonstration of the extent to which predators can influence lower trophic levels. □

Received 12 October 1989; accepted 11 January 1990.

- Hrbacek, J. *Cesk. Akad. Ved. Rada. Mat. Prirod. Ved.* **72**, 1–116 (1962).
- Brooks, J. L. & Dodson, S. I. *Science* **150**, 28–35 (1965).
- Paine, R. T. *J. Animal Ecol.* **49**, 667–685 (1980).
- A. Sih, Crowley, P., McPeck, M., Petranka, J. & Strohmeier, K. *A. Rev. ecol. Syst.* **16**, 269–311 (1985).
- Kerfoot, W. C. & Sih, A. (eds) *Predation: Direct and Indirect Impacts on Aquatic Communities* (University Press of New England, Hanover, New Hampshire, 1987).
- Menge, B. A. & Sutherland, J. P. *Am. Nat.* **130**, 730–757 (1987).
- Carpenter, S. R. (ed) *Complex Interactions in Lake Communities* (Springer, New York, 1988).
- Carpenter, S. R., Kitchell, J. F. & Hodgson, J. R. *BioScience* **35**, 634–639 (1985).
- Carpenter, S. R. & Kitchell, J. F. *BioScience* **38**, 764–769 (1988).
- Shapiro, J. & Wright, D. I. *Freshwater Biol.* **14**, 371–383 (1984).
- Lynch, M. *Limnol. Oceanogr.* **24**, 253–272 (1979).
- Bergquist, A. M. & Carpenter, S. R. *Ecology* **67**, 1351–1360 (1986).
- Threlkeld, S. T. & Drenner, R. W. *Limnol. Oceanogr.* **32**, 1331–1341 (1987).
- Vanni, M. J. *Ecol. Monogr.* **57**, 61–88 (1987).
- Neill, W. E. in *Complex Interactions in Lake Communities* (ed. Carpenter, S. R.) (Springer, New York, 1988).
- Brock, T. D. *A Eutrophic Lake: Lake Mendota, Wisconsin* (Springer, New York, 1985).
- Vanni, M. J. & Temte, J. *Limnol. Oceanogr.* (in the press).
- Luecke, C., Vanni, M. J., Magnuson, J. J., Kitchell, J. F. & Jacobson, P. J. *Limnol. Oceanogr.* (submitted).
- Pedros-Allo, C. & Brock, T. D. *Freshwater Biol.* **15**, 89–94 (1985).
- Lampert, W., Fleckner, W., Rai, H. & Taylor, B. E. *Limnol. Oceanogr.* **31**, 478–490 (1986).
- Knoechel, R. & Holtby, L. B. *Limnol. Oceanogr.* **31**, 1–16 (1986).
- John, K. R. & Hasler, A. D. *Limnol. Oceanogr.* **1**, 176–194 (1956).
- Wetzel, R. G. & Likens, G. E. *Limnological Analyses* (Saunders, Philadelphia, 1979).
- Lynch, M., Weider, L. J. & Lampert, W. *Limnol. Oceanogr.* **31**, 17–31 (1986).

ACKNOWLEDGEMENTS. We thank our colleagues at the Wisconsin Department of Natural Resources, B. Johnson, S. Gilbert, and especially R. C. Lathrop, who collected some of the nutrient samples and discussed with us events occurring in Lake Mendota from 1982–1985. Reviews by S. R. Carpenter and T. M. Frost improved the manuscript. This research was supported by a grant from the Wisconsin Department of Natural Resources (J.F.K. and J.J.M.).

Precocious flowering and seeding behaviour in tissue-cultured bamboos

R. S. Nadgouda, V. A. Parasharami
& A. F. Mascarenhas

Division of Biochemical Sciences, National Chemical Laboratory,
Pune-411 008, India

BAMBOO flowers only once during its lifetime, dying at the end of its first fruiting season. This monocarpic flowering is intriguing not only in that it occurs after a lapse of 12 to 120 years, but because it is 'gregarious', local populations of bamboo flowering together and then dying. New bamboo plants are produced either by vegetative subdivision or from seed. Breeding of bamboo, however, has proved to be extremely difficult: seed production depends on unpredictable circumstances and events¹, and the basis of gregarious flowering, and the causes of death and flowering, are not known. Flowering *in vitro* has previously been studied by culturing explants of stem tips, mature stems, roots, petioles, leaves, inflorescences, flowers and so on^{2–5}. Although bamboo plantlets have been formed by means of organogenesis and embryogenesis^{6–8}, *in vitro* flowering has not previously been reported for bamboo. We now report on an *in vitro* system in which we could consistently induce flowering in the two species of bamboo *Bambusa arundinacea* Willd and *Dendrocalamus brandisii* Kurz. Inflorescence explants containing a panicle of spikelets gave rise to several viable inflorescences on subculture; fertile seeds were also produced. Further refinements to this system could lead to the introduction of breeding programmes to improve bamboo, and to the production of perennial seeds for bamboo, as well as to a better understanding of the physiology underlying flowering behaviour in bamboo.

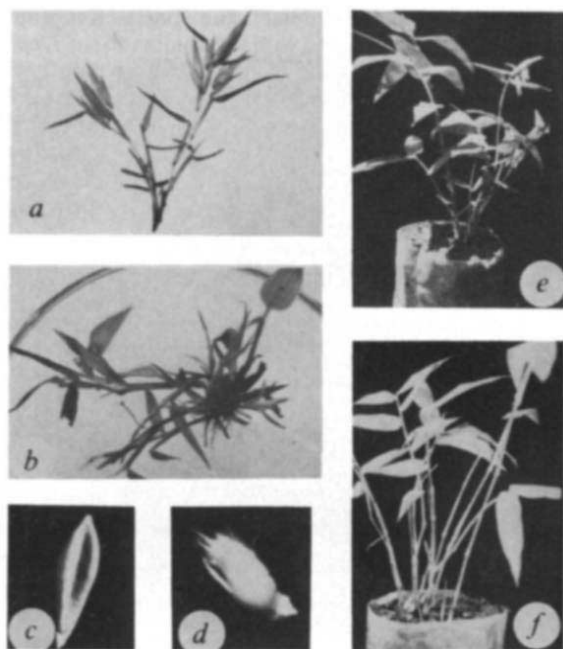


FIG. 1 *a* and *b*, Inflorescence cultures: *B. arundinacea* (*a*); *D. brandisii* (*b*). *c* and *d*, Seeds ($\times 30$): *B. arundinacea* (*c*); *D. brandisii* (*d*). *e*, *B. arundinacea* in soil showing inflorescence. *f*, *B. arundinacea* in soil, grown after harvest of seeds.

We germinated seeds of the two bamboo species *D. brandisii* and *B. arundinacea* in the dark on modified White's medium⁹ adjusted to pH 5.8 and containing 2% sucrose and 0.4% agar. After 8 days, when seeds germinated, they were transferred to light (500 lx) at $28 \pm 2^\circ\text{C}$ where they grew to a height of 5–6 cm.

At this stage, we excised 3–4-cm segments containing the coleoptile regions with the growing points, and transferred them to Murashige and Skoog's (MS)¹⁰ liquid medium containing 2% sucrose, and incubated them at 28°C under light (500 lx) on a rotary shaker at 120 r.p.m. Under these conditions, a clump of shoots developed. To study growth responses, we then excised individual shoots and subcultured them on MS medium supplemented either with coconut milk (CM) or other growth regulators singly or in different combinations and concentrations. Each experiment was performed using 10 replicates. MS basal medium supplemented with 2% sucrose and 0.5 p.p.m. benzyl-amino-purine (BAP) and 5% CM gave the best response, resulting in further growth of 15–20 shoots per culture vessel.

After three consecutive subcultures on this medium, 70% of *B. arundinacea* cultures and 40% of *D. brandisii* cultures developed panicles of normal spikelets (Fig. 1*a, b*).

Flowers of both the species were about 0.5 to 1 cm and consisted of an outer palea and inner lemma. In both species, stamens and pistils were normal.

Out of a cluster of 15–20 shoots present in each culture vessel, 60% gave rise to inflorescences in both species. The remaining shoots appeared to be vegetative. We then dissected individual vegetative or inflorescence shoots, and cultured and maintained them separately. We rooted separated vegetative shoots of both the species *in vitro* by treatment with 0.5 p.p.m. indole butyric acid for 24 h, and then transferred them to White's liquid medium for 15 days. The rooted plants were then transferred to soil (Fig. 1*e, f*).

Inflorescence explants containing the panicle of spikelets in both species gave rise to several inflorescences on subculture to the same medium, thereby enabling maintenance of an inflorescence culture. We observed seed set in cultures of both species

grown (Fig. 1*c, d*) *in vitro* (Fig. 2*a, b*). About 50 normal seeds were obtained from each culture of *B. arundinacea*, whereas each *D. brandisii* culture produced about five.

We are now attempting to find the most favourable sets of environmental and nutritional conditions for flower induction and seed formation *in vitro*.

The observations reported here are novel, and with further experiments should lead to a better understanding of: (1) the physiological and molecular events underlying the shift from the stable vegetative state to the monocarpic floral state; (2) the specific roles of cytokinins in inducing precocious flower development; and (3) the causes of gregarious flowering and death of bamboo plants.

Bamboo has an important role in the world's industrial and domestic economics; our findings have vast potential in this respect, in that they could pave the way for the breeding of improved bamboo, and for the production of interspecific-intergeneric hybrids, as well as for the provision of a source of perennial bamboo seed.

The results described here have so far been reproduced three times with the two species, and recently with *D. strictus*. □

Received 20 December 1989; accepted 1 February 1990.

1. Janzen, D. H. *A. Rev. ecol. Syst.* **7**, 347–391 (1976).
2. Pierik, R. L. M. *Naturwissenschaften* **53**, 45 (1966).
3. Nitsch, C. & Nitsch, J. P. *Planta* **72**, 371–384 (1967).
4. Paulet, P. & Nitsch, J. P. *Annls Physiol. vég Paris* **6**, 333–345 (1964).
5. Chang, W.-c. & Hsing, Y.-i. *Nature* **284**, 341–342 (1980).
6. Nadgir, A. L., Phadke, C. H., Gupta, P. K., Parasharami, M. & Mascarenhas, A. F. *Silvae Genet.* **33**, 219–223 (1984).
7. Rao, V., Rao, I. R. & Narang, V. *Pl. Cell Reps* **4**, 191–194 (1985).
8. Hassan, A. A. & Debergh, P. *Pl. Cell Tiss. Organ Culture* **10**, 73–77 (1987).
9. White, P. R. *The Cultivation of Animal and Plant Cells* 2nd edn (Ronald, New York, 1963).
10. Murashige, T. & Skoog, F. *Physiol. Pl.* **15**, 473–479 (1962).

ACKNOWLEDGEMENTS. We thank Drs S. Mantell and Jennet Blake at Wye College, University of London for their helpful suggestions, P. K. Salma and E. Muralidharan for technical assistance, and Bridget Cue for preparing the manuscript.

Dual ion-channel regulation by cyclic GMP and cyclic GMP-dependent protein kinase

Douglas B. Light*, Jackie D. Corbin† & Bruce A. Stanton

Department of Physiology, Dartmouth Medical School, Hanover, New Hampshire, USA

† Department of Molecular Physiology and Biophysics, Vanderbilt University Medical School, Nashville, Tennessee 37232, USA

ATRIAL natriuretic peptide, acting through its second messenger guanosine 3',5'-cyclic monophosphate (cGMP), suppresses Na^+ absorption across the renal inner-medullary collecting duct and increases urinary Na^+ excretion. Patch clamp studies show that cGMP reduces Na^+ absorption by inhibiting an amiloride-sensitive cation channel in the apical membrane. We have now examined, using the patch clamp technique, the molecular mechanisms of cGMP inhibition. Cyclic GMP directly and specifically reduced the probability of a single channel being open (open probability, P_o) by 39% (inhibition constant, $K_i = 7.6 \times 10^{-7} \text{ M}$) by a phosphorylation-independent mechanism. Cyclic GMP also inhibited the channel by activating cGMP-dependent protein kinase (cGMP-kinase). Exogenous cGMP-kinase completely inhibited the channel by a phosphorylation-dependent mechanism. Activation of a pertussis toxin-sensitive G protein by GTP- γ -S blocked cGMP-kinase inhibition of the channel. By contrast, cGMP-kinase inhibi-

* Present address: Department of Biology, Ripon College, Ripon, Wisconsin 54971, USA.

tion of P_o was completely reversed by GTP- γ -S. Taken together with the results of a previous study showing that a G protein activates the cation channel, these data indicate that cGMP-kinase and a G protein sequentially regulate the cation channel. Our results show that atrial natriuretic peptide, acting through cGMP, inhibits Na^+ absorption across the inner-medullary collecting duct by a dual mechanism, and that cGMP-kinase inhibits the channel by a pathway involving a G protein.

Atrial natriuretic peptide (ANP), a circulating hormone released from the atria of mammalian hearts in response to an increase in extracellular fluid volume, increases urinary excretion of Na^+ partly by inhibiting electrogenic Na^+ absorption across the renal inner-medullary collecting duct (IMCD)¹⁻⁴. ANP, acting through cGMP, reduces Na^+ absorption by inhibiting an amiloride-sensitive ($K_i < 10^{-7}$ M) cation channel (28 pS, $P_{\text{Na}}:P_{\text{Cl}} = 13:1$) in the apical membrane of rat IMCD cells⁴⁻⁶. The objective of the study described here was to determine the molecular mechanisms of this inhibition.

Cyclic GMP in the bath solution of inside-out patches of the apical membrane reversibly inhibited the channel. Maximum inhibition was achieved with 10^{-4} M cGMP (39% reduction in P_o), and the concentration for half-maximal inhibition was 7.6×10^{-7} M (Table 1). Neither adenosine 3',5'-cyclic monophosphate (cyclic AMP), cytidine 3',5'-cyclic monophosphate (cCMP), thymidine 3',5'-cyclic monophosphate (cTMP) nor uridine 3',5'-cyclic monophosphate (cUMP) inhibited the channel (Table 1). We performed several experiments to determine

TABLE 1 Effects of cyclic nucleotides, Mg^{2+} and H8 on the cation channel in excised inside-out patches of the apical membrane

Compound	Concentration (M)	P_o (control)	P_o (experimental)	% Change
cGMP ($n=3$)	10^{-6}	0.72 ± 0.08	$0.58 \pm 0.13^*$	19.4
cGMP ($n=10$)	10^{-5}	0.70 ± 0.04	$0.50 \pm 0.07^{**}$	28.6
cGMP ($n=14$)	10^{-4}	0.76 ± 0.04	$0.46 \pm 0.06^{**}$	39.5
cGMP ($n=3$)	10^{-3}	0.85 ± 0.05	$0.55 \pm 0.01^*$	35.3
cAMP ($n=8$)	10^{-3}	0.46 ± 0.10	0.43 ± 0.07	6.5
cCMP ($n=3$)	10^{-4}	0.78 ± 0.10	0.76 ± 0.10	2.6
cTMP ($n=3$)	10^{-4}	0.58 ± 0.15	0.60 ± 0.19	3.4
cUMP ($n=3$)	10^{-4}	0.39 ± 0.19	0.35 ± 0.17	10.3
0 M Mg^{2+} solution ($n=4$)	0	0.80 ± 0.07	0.76 ± 0.12	5.0
cGMP in 0 M Mg^{2+} solution ($n=4$)	10^{-4}	0.76 ± 0.12	$0.60 \pm 0.12^*$	22.2
H8 ($n=7$)	10^{-6}	0.79 ± 0.11	0.76 ± 0.12	3.8
cGMP with H8 ($n=7$)	10^{-4}	0.76 ± 0.12	$0.49 \pm 0.13^*$	36.0

Cyclic nucleotides and H8 (10^{-6} M) were added to the bath solution in paired experiments. The solutions did not contain ATP. The 0 M Mg^{2+} solution contained 1 mM EDTA. Data are mean $\pm$ s.e.m. (n =number of patches studied). Statistical significance was determined by the paired Student's t -test. *, $P < 0.05$; **, $P < 0.01$.

whether the specific inhibition of the channel by cGMP was mediated by a phosphorylation-independent mechanism (Table 1). Cyclic GMP (10^{-4} M) decreased P_o even in the absence of Mg^{2+} , a cofactor required for kinase activity⁷, and in the presence of the protein kinase inhibitor N -[2-(methylamino)-ethyl]-5-isoquinolinesulphonamide dihydrochloride (H8) (ref. 8). H8 (10^{-6} M) alone, or removal of Mg^{2+} from the bath solution, had no effect on the P_o (Table 1). Taken together, these results show that cGMP inhibits the cation channel directly and specifically by a mechanism that does not involve phosphorylation.

ANP, which elevates cGMP in IMCD cells, completely inhibits the cation channel in cell-attached patches⁴. Because the phosphorylation-independent inhibition by cGMP in inside-out patches reduced the P_o by only 39%, we next examined whether complete inhibition of the channel by cGMP requires active cGMP-kinase. Figure 1 illustrates a representative experiment. Activation of cGMP-kinase requires both cGMP and ATP¹¹. Cyclic GMP (10^{-5} M) and ATP (10^{-5} M) together

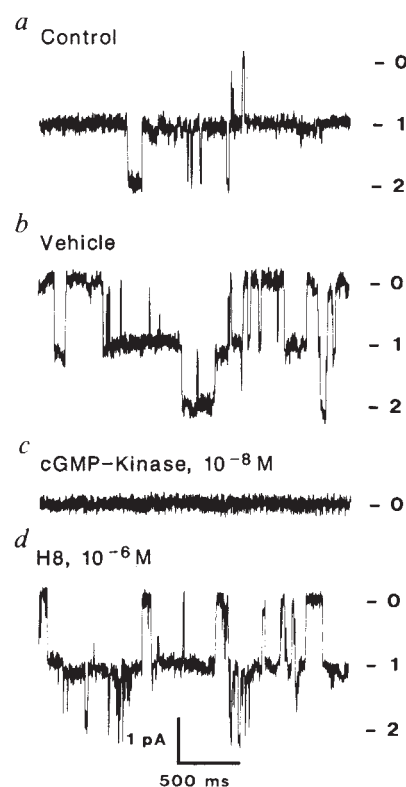


FIG. 1 Current records from an excised inside-out patch of the apical membrane illustrating phosphorylation-dependent inactivation of the cation channel by cGMP-kinase. *a*, Control. Two channels were present in the patch. *b*, Thirty seconds after the addition of vehicle (10^{-5} M cGMP, 10^{-5} M ATP) to the bath. Cyclic GMP and ATP are cofactors required for cGMP-kinase activity⁷. In 10 experiments, the P_o decreased from 0.68 ± 0.06 to 0.52 ± 0.06 ($P < 0.01$). *c*, Two minutes after the addition of cGMP-kinase (10^{-8} M) to the bath (with vehicle). The P_o decreased to 0.00 ($P < 0.05$). Cyclic GMP-kinase inhibited the channel in every experiment. *d*, Four minutes after the addition of H8 (10^{-6} M) to the bath (in the presence of cGMP-kinase, cGMP and ATP; $n=4$). The P_o increased from 0 to 0.46 ± 0.10 ($P < 0.05$). Patch clamp experiments were conducted at 23 °C on rat IMCD cells in primary culture as described previously^{4,5}. Single-channel ion currents were measured with a current-to-voltage converter, low-passed filtered at 100 Hz, and digitized at 1 kHz with an IBM AT computer. Data were recorded for at least three 20-s periods during the control and experimental periods (with a 30-s delay between periods). All experimental manoeuvres were reversible unless stated otherwise. The P_o was estimated from the total time spent in the open state divided by the total time of the record as described^{4,5}. All experiments were conducted in a paired fashion (that is, each patch served as its own control). Data are reported as means $\pm$ s.e.m. Statistical significance was determined by the paired Student's t -test. Cyclic GMP-kinase was extracted from bovine lung and purified to homogeneity (>99%) using an affinity column as described^{27,28}. Pipette (cell-attached and inside-out patches) and bath (cell-attached patches) solutions (in mM): NaCl, 140; KCl, 5; CaCl_2 , 1; MgCl_2 , 1; HEPES buffer, 10; pH 7.4. Bath (inside-out patches) solution: NaCl, 5; KCl, 140; CaCl_2 , 1; MgCl_2 , 1; HEPES buffer, 10; pH 7.4. The 0 Mg^{2+} solution also contained 1 mM EDTA. The command voltage in all records was -60 mV. Voltage refers to the cytoplasmic side of the membrane referenced to the interior of the patch pipette. The numbers to the right of each current record indicate the number of open channels.

decreased the P_o from 0.68 ± 0.06 to 0.52 ± 0.06 ($n=11$, $P < 0.01$; Fig. 1*b*). This inhibition was similar to that produced by cGMP alone. Active exogenous cGMP-kinase (10^{-7} or 10^{-8} M) reduced the P_o by 96.1% ($n=10$; $P < 0.001$; Fig. 1*c*). Cyclic GMP-kinase inhibition of the channel had an absolute requirement for both cGMP and ATP. Cyclic GMP-kinase had no effect when the bath solution contained either cGMP or ATP alone. Heat-inactivated cGMP-kinase (100 °C for 30 min), in the presence of cGMP and ATP, was without effect. Removal of Mg^{2+} from the bath solution (containing 1 mM EDTA) prevented the

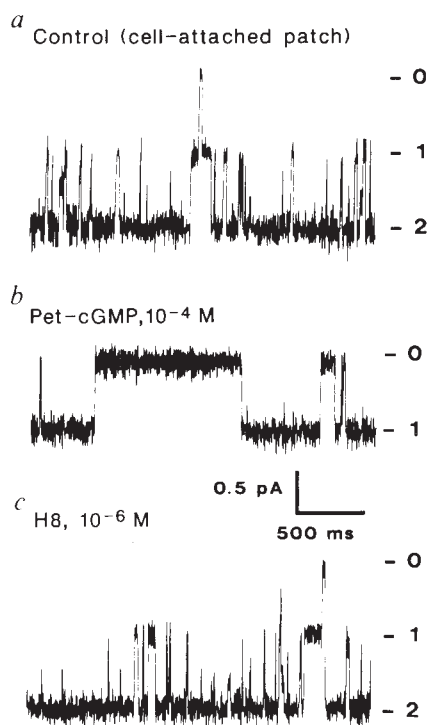


FIG. 2 Current records from a cell-attached patch of the apical membrane illustrating the presence of endogenous cGMP-kinase in IMCD cells. *a*, Control. Two channels were present in the patch and both were open most of the time. *b*, Two minutes after the addition of Pet-cGMP (10^{-4} M), synthesized as described¹⁰. Simultaneous channel openings were not observed. P_o decreased sharply. *c*, Two minutes after the addition of H8 (10^{-6} M) to the bath in the presence of Pet-cGMP. The single-channel P_o increased. In all records, the command voltage (V_c) was +60 mV; the voltage across the patch of membrane was the sum of V_c and the resting membrane voltage (~ -60 mV). The numbers to the right of each current record indicate the number of open channels.

inhibition by cGMP-kinase, and H8 (10^{-6} M) reversed the inhibitory effect of cGMP-kinase (Fig. 1d). Together these results show that cGMP-kinase inhibits the channel by an apparent phosphorylation-dependent mechanism.

The catalytic subunit of cAMP-dependent protein kinase (cAMP-kinase, 6.3×10^{-8} M; Mg-ATP, 10^{-3} M), had no effect on the channel ($n = 3$, 5-min exposure⁵). Because of the similarity in substrate specificity of cGMP-kinase and cAMP-kinase⁷, we cannot rule out the possibility that a higher concentration of cAMP-kinase would inhibit the channel.

The kidney is a rich source of cGMP-kinase⁹. To determine whether cGMP-kinase is present in intact IMCD cells in culture and, when activated by cGMP, inhibits the cation channel, we added 1-*N*²-phenylethano-cGMP (Pet-cGMP) to the bath solution of cell-attached patches. Pet-cGMP is a membrane-soluble poorly hydrolysable analogue of cGMP that potently and specifically activates cGMP-kinase both *in vitro* and in intact tissue¹⁰. Pet-cGMP (10^{-4} M) decreased the P_o from 0.78 ± 0.07 to 0.29 ± 0.09 ($n = 4$, $P < 0.001$; Fig. 2a and b). H8 partially reversed the action of Pet-cGMP: the P_o increased to 0.63 ± 0.23 (Fig. 2c). These observations indicate that endogenous cGMP-kinase, when activated by cGMP, inhibits the cation channel.

The α_{i-3} subunit of the pertussis toxin-sensitive G protein, G_i , activates the cation channel in IMCD cells¹¹. We therefore examined whether cGMP-kinase inhibits the channel by a pathway involving G_i . GTP- γ -S activates G proteins and also prevents kinase phosphorylation of some G proteins¹²⁻¹⁵. To determine whether GTP- γ -S could block the inhibitory effect of cGMP-kinase, we first added GTP- γ -S to the bath solution of inside-out patches and then added cGMP-kinase. Figure 3 illustrates a representative experiment. GTP- γ -S (10^{-4} M) increased the P_o from 0.43 ± 0.12 to 0.68 ± 0.29 ($n = 5$, $P < 0.01$). Addition of the vehicle for cGMP-kinase (10^{-4} M cGMP, 10^{-4} M ATP) to the bath solution, in the presence of GTP- γ -S, reduced the P_o to 0.42 ± 0.12 ($P < 0.05$). Therefore cGMP inhibits the cation channel even when G_i is activated by GTP- γ -S. Cyclic GMP-kinase (10^{-7} M) had no effect on P_o when the G protein was activated with GTP- γ -S (Fig. 3d). The order of the experimental procedures was then reversed to determine whether GTP- γ -S could relieve kinase inhibition. Active cGMP-kinase was first added to the bath; as shown above, cGMP-kinase closed the

channel. GTP- γ -S completely reversed cGMP-kinase inhibition ($n = 4$). These experiments indicate that cGMP-kinase and G_i might regulate the channel by a sequential pathway. Cyclic GMP-kinase could modulate the ability of G_i to activate the cation channel; but, additional experiments are required to elucidate completely the molecular interactions among cGMP-kinase, G_i and the cation channel.

Our experiments show that cGMP, a second messenger of ANP, inhibits a cation channel in the apical membrane of IMCD cells directly, most probably by interacting with an allosteric modifier site on the channel, and by a phosphorylation-dependent mechanism involving cGMP-kinase. Because the K_m of cGMP for cGMP-kinase ($0.7\text{--}1.6 \times 10^{-7}$ M) is about 5–10-fold lower than the apparent K_i of cGMP for the cation channel (7.6×10^{-7} M), it could be argued that cGMP-kinase would dominate this regulatory system. The absolute contribution of each pathway, however, cannot yet be determined because of uncertainties about measurements of intracellular cGMP, possible compartmentalization of cGMP, and difficulties in determining the kinetic reactions in intact cells at 37 °C.

Cyclic GMP regulates nonselective cation channels in renal IMCD cells, in rod and cone photoreceptors and in olfactory cells¹⁶⁻²². The Michaelis constant K_m for cGMP activation is 10–20-fold higher in rod and cone cells and two-fold higher in olfactory cells than the K_i for the renal cation channel¹⁶⁻²². The cation channel in rod and cone photoreceptor cells, like the renal channel, has a strong selectivity for cGMP over other cyclic nucleotides^{16,18,21}. By contrast, the channel in olfactory cells is activated by cGMP and cAMP²². These studies indicate that there is a family of cGMP-gated cation channels. Because cGMP inhibited the renal cation channel by a phosphorylation-independent pathway, and because cGMP does not increase the activity of G proteins we suggest that, by analogy with the rod and cone photoreceptor, cGMP binds to an allosteric modifier site on the renal cation channel¹⁶⁻²¹.

The effect of cGMP-kinase on the cation channel in cone, rod and olfactory cells has not been reported. Cyclic GMP-kinase regulates conductive pathways in other cells. In snail neurons it potentiates the serotonin-induced macroscopic Ca^{2+} current^{23,24} and it increases the input resistance in neurons from the mammalian motor cortex²⁵. The interaction between cGMP-

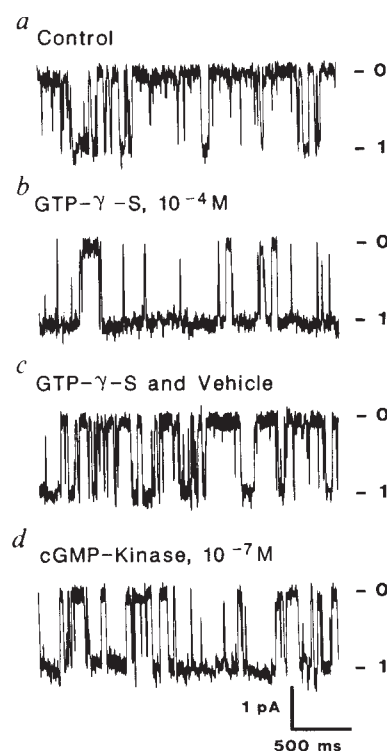


FIG. 3 Current records from an inside-out patch of the apical membrane illustrating the involvement of a G protein in the action of cGMP-kinase. *a*, Control. One channel was present in the patch. *b*, One minute after the addition of GTP- γ -S (10^{-4} M) to the bath. The P_o increased. *c*, One minute after the addition of vehicle (10^{-4} M cGMP, 10^{-4} M ATP) in the presence of GTP- γ -S. P_o decreased. *d*, Two minutes after the addition of cGMP-kinase (10^{-7} M) in the presence of cGMP, ATP and GTP- γ -S. P_o did not change when cGMP-kinase was added. The voltage across the patch was -60 mV. The numbers to the right of each current record indicate the numbers of open channels.

kinase and a G protein in regulating these conductances has not been reported²⁶.

Inhibition of the cation channel by two cGMP-dependent mechanisms would produce both a short-lived and a sustained reduction in Na^+ absorption by the IMCD. Cyclic GMP inhibition of the channel in inside-out patches was immediately reversible (<2 s). By contrast, reversal of cGMP-kinase inhibition, after administration of H8 or removal of kinase from the bath of inside-out patches, generally required 2–3 min, although in some experiments the inhibition did not reverse (~ 10 min). This apparent irreversibility could result from a low phosphatase activity in some membrane patches. Accordingly, we suggest that *in vivo*, the allosteric effect of cGMP would terminate rapidly (<2 s) after a decrease in plasma ANP and, thereby, intracellular cGMP, whereas the phosphorylation-dependent reduction in Na^+ absorption would persist for at least several minutes. $\square$

Received 2 October 1989; accepted 23 January 1990.

- DeBoid, A. J. *Fedn Proc.* **45**, 2081–2085 (1986).
- Maack, T. & Klienert, H. D. *Biochem. Pharmacol.* **35**, 2057–2064 (1986).
- Zeidel, M. L., Kikeri, D., Silva, P., Burrows, M. & Brenner, B. M. *J. clin. Invest.* **82**, 1067–1074 (1988).
- Light, D. B., Schwiebert, E. M., Karlson, K. H. & Stanton, B. A. *Science* **243**, 383–385 (1989).
- Light, D. B., McCann, F. V., Keller, T. M. & Stanton, B. A. *Am. J. Physiol.* **255**, F278–F286 (1988).
- Stanton, B. A. *Am. J. Physiol.* **256**, F862–F868 (1989).
- Beebe, S. J. & Corbin, J. D. *The Enzymes* Vol. 17 (eds Krebs, E. G. & Boyer, P. D.) 43–111 (Academic, Orlando & London, 1986).
- Hidaka, H., Inagaki, M., Kawamoto, S. & Saaki, Y. *Biochemistry* **23**, 5036–5041 (1984).
- Walter, U. *Eur. J. Biochem.* **118**, 339–346 (1981).
- Francis, S. H., Noblett, B. D., Todd, B. W., Wells, J. N. & Corbin, J. D. *Molec. Pharmacol.* **34**, 506–517 (1988).
- Light, D. B., Ausiello, D. A. & Stanton, B. A. *J. clin. Invest.* **84**, 352–356 (1989).
- Katada, T., Gilman, A. G., Watanabe, Y., Bauer, K. H. & Jakobs, K. H. *Eur. J. Biochem.* **151**, 431–437 (1985).
- Zick, Y., Sagi-Eisenberg, R., Pines, M., Gierschik, P. & Spiegel, A. M. *Proc. natn. Acad. Sci. U.S.A.* **83**, 9294–9297 (1986).

- Yada, Y., Nagao, S., Okano, Y. & Nozawa, Y. *FEBS Lett.* **242**, 368–372 (1989).
- Nagata, K., Nagao, S. & Nozawa, Y. *Biochem. biophys. Res. Commun.* **160**, 235–242 (1989).
- Fesenko, E. E., Kolesnikov, S. S. & Lyubarsky, A. L. *Nature* **313**, 310–313 (1985).
- Fesenko, E. E., Kolesnikov, S. S. & Lyubarsky, A. L. *Biochem. biophys. Acta* **856**, 661–671 (1986).
- Tanaka, J. C., Furman, R. E., Cobbs, W. H. & Mueller, P. *Proc. natn. Acad. Sci. U.S.A.* **84**, 724–728 (1987).
- Cook, N. J., Hanke, W. & Kaupp, U. B. *Proc. natn. Acad. Sci. U.S.A.* **84**, 585–589 (1987).
- Haynes, L. W., Kay, A. R. & Yau, K.-W. *Nature* **321**, 66–70 (1986).
- Haynes, L. & Yau, K. W. *Nature* **317**, 61–64 (1985).
- Nakamura, T. & Gold, G. H. *Nature* **325**, 442–444 (1987).
- Paupardin-Tritsch, D., Hammond, C., Gerschenfeld, H. M., Narin, A. C. & Greengard, P. *Nature* **323**, 812–814 (1986).
- Hemmings, H. C., Nairn, A. C., McGuinness, T. L., Huganir, R. L. & Greengard, P. *FASEB J.* **3**, 1583–1592 (1989).
- Woody, C. D., Bartfal, T., Gruen, E. & Nairn, A. C. *Brain Res.* **386**, 379–385 (1986).
- Levitin, I. B. A. *Rev. Neurosci.* **11**, 119–136 (1988).
- Wolf, L., Francis, S. H., Landiss, L. R. & Corbin, J. D. *J. biol. Chem.* **262**, 16906–16913 (1987).
- Corbin, J. D., Ogried, D., Miller, J. P., Suva, R. H., Jastorff, B. & Doskeland, S. O. *J. biol. Chem.* **261**, 1208–1214 (1986).

ACKNOWLEDGEMENTS. We thank Dr Sharron Francis for advice, Erik Schwiebert for analysing current records, Dr Frances McCann for the computer program for analysing records and Katherine Karlson for technical assistance. This work was supported by the NIH, the American Heart Association and Wyeth-Ayerst (B.S.).

Identification and characterization of a novel member of the nerve growth factor/brain-derived neurotrophic factor family

Andreas Hohn, Joachim Leibrock, Karen Bailey & Yves-Alain Barde

Max-Planck-Institut für Psychiatrie, Department of Neurochemistry, 8033 Planegg-Martinsried, FRG

THE survival and functional maintenance of vertebrate neurons critically depends on the availability of specific neurotrophic factors^{1–3}. So far, only two such factors, nerve growth factor (NGF) and brain-derived neurotrophic factor (BDNF) have been characterized and shown to have the typical features of secretory proteins. This characterization has been possible because of the extraordinarily large quantities of NGF in some adult tissues⁴, and the virtually unlimited availability of brain tissue from which BDNF was isolated⁵. Both NGF and BDNF promote the survival of distinct neuronal populations *in vivo* and are related in their primary structure⁶, suggesting that they are members of a gene family. Although there is little doubt about the existence of other such proteins, their low abundance has rendered their identification and characterization difficult. Taking advantage of sequence identities between NGF and BDNF, we have now identified a third member of this family, which we name neurotrophin-3. Both the tissue distribution of the messenger RNA and the neuronal specificity of this secretory protein differ from those of NGF and BDNF. Alignment of the sequences of the three proteins reveals a remarkable number of amino acid identities, including all cysteine residues. This alignment also delineates four variable domains, each of 7–11 amino acids, indicating structural elements presumably involved in the neuronal specificity of these proteins.

A comparison of the mouse NGF and BDNF amino-acid sequences revealed two stretches of six amino acids (underlined in Fig. 1) that are conserved in each of the proteins. We synthesised two oligonucleotide primer pools based on these two sequences, and used them to prime the amplification of a mouse genomic template using the polymerase chain reaction (PCR)⁷. The resulting DNA fragments were then digested using restriction enzymes specific for NGF or BDNF sequences. We sequenced an uncleaved DNA fragment, showing that it did not correspond to either NGF or BDNF genes. This sequence was then used to synthesize two specific sense (or 5') primers for further PCRs on a complementary DNA template prepared by reverse transcription of RNA extracted from mouse brain, muscle and liver (Fig. 1). These three tissues gave amplified

FIG. 1 Genomic sequence and deduced amino-acid sequence of NT-3. The amino-acid sequence starts with the first ATG codon found after three in-frame stop codons. The underlined sequences indicate the location of the primers used in the first round of PCR. The only *N*-glycosylation consensus sequence is doubly underlined, and the arrow indicates the putative start of processed mature NT-3.

METHODS. Two oligonucleotide primer pools were synthesized on the basis of two amino-acid sequences conserved among BDNF and all known NGFs⁶. The sense (5') primer (GGGGATCCGCG GGI TGY MGI GGI ATH GA; primer 1; IUPAC nomenclature; I, inosine) includes restriction sites for *Bam*HI and *Sac*II and the antisense (3') primer (TCGAATTCTAG AT ICK IAT RAA ICK CCA; primer 2) includes *Eco*RI and *Xba*I sites. PCR was performed with 1 µg of mouse genomic DNA using a Perkin-Elmer Cetus thermal cycler and *Taq* polymerase (GeneAmpTM). After four cycles with an annealing temperature of 45 °C, the remaining 36 cycles were performed with annealings at 49 °C. The resulting amplified DNA products corresponding to the expected size (137 base pairs (bp)) were eluted from an acrylamide gel, reamplified, and digested with *Hind*III and *Apa*I that cleave mouse NGF and mouse BDNF, respectively. The uncleaved DNA was eluted and asymmetrically amplified¹⁶, and sequenced¹⁷ using primers 1 and 2. The sequence thus obtained was used to synthesize primers corresponding to nucleotides 808–822 (primer 3) and 824–844 (primer 4), to which *Sal*I and *Pst*I sites were added. RNA was extracted from adult mouse brain, liver and muscle¹⁸, and reverse-transcribed using the antisense (5') primer CGGATCCGAATTCTGCAG(T)₁₂V (primer 5, designed to match 3' poly(A) tails and containing cloning sites for *Bam*HI, *Eco*RI and *Pst*I⁶). These complementary DNAs were first PCR-amplified using primers 3 and 5, and re-amplified using primers 4 and 5. A Southern-blot analysis was performed with the re-amplified products, and hybridized with a ³²P-end-labelled oligonucleotide corresponding to nucleotides 879–893. The DNA fragments thus identified were cloned into a Bluescript SK⁺ vector (Stratagene), and the longest insert obtained (460 bp) was used to screen an EMBL3 mouse genomic library (Clontech). The DNA of one of the two positive clones thus identified was digested with various restriction enzymes. The restriction fragments were probed with the 460-bp insert: a 770-bp *Hind*III fragment,

products of similar sizes that were cloned and used to screen a mouse genomic library. One of the positive genomic clones thus identified was sequenced (Fig. 1), showing an open reading frame predicted to encode a protein of 258 amino acids. We named this protein neurotrophin-3 (NT-3).

In every respect, the overall structure of NT-3 resembles that of NGF and BDNF: a putative signal sequence of 18 amino acids (showing five and nine amino-acid identities with BDNF and NGF, respectively) is followed by a prosequence of 121 amino acids. Such pro-sequences, presumably involved in the folding and correct formation of the disulphide bridges of these proteins⁸, have also been found in mouse NGF (103 amino acids) and mouse BDNF (112 amino acids). A single potential *N*-glycosylation site is located nine amino acids (eight in BDNF and NGF) before the putative cleavage site, characterized by a cluster of basic residues, for the generation of mature NT-3 (Fig. 1, arrow). Mature NT-3 is predicted to consist of 119 amino acids (relative molecular mass of 13,625, pI 9.3). Comparison of mature mouse NGF, BDNF and NT-3 reveals 54 amino-acid identities (Fig. 2). All six cysteine residues, which in NGF and BDNF are involved in the formation of disulphide bridges^{6,9}, are among the conserved residues (Fig. 2, arrows). Remarkably, the addition of the NT-3 sequence to those of NGF and BDNF results in the loss of only seven amino-acid identities (there are 61 amino-acid identities between mouse NGF and BDNF). Therefore, almost 50% of the structure is conserved among the three proteins. In addition, comparison of the three sequences reveals four variable domains, each of 7–11 amino acids (numbered V1 to V4 in Fig. 2); these are presumably involved in the neuronal specificity of these proteins.

To study the expression of the NT-3 gene, we extracted RNA

		GGAAACCCCG CCGCTATAA ATAACAGGGA GGAGTTTACC		40
TCATTGTGTA GACTGGTGGT ACCCTCTCCT		CAGTCTCAGA TTGTCCAGCT CCGTGGCTGA TGCCAGAAATG		110
ACAGGGGGCTC TTCCGGCCAG CTGTAAACAC		CTGTGTITTC TTTTTCAGA TCTTACAGGT GAACAAGGTG		180
1		10		20
Met Ser Ile Leu Phe Tyr Val Ile Phe	Leu Ala Tyr Leu Arg Gly Ile Gln Gly Asn Ser			240
ATG TCC ATC TTG TTT TAT GTG ATA TTT	CTT GCT TAT CTC CGT GGC ATC CAA GGC AAC AGC			40
30		40		300
Met Asp Gln Arg Ser Leu Pro Glu Asp	Ser Leu Asn Ser Leu Ile Ile Lys Leu Ile Gln			360
ATG GAT CAA AGG AGT TTG CCG GAA GAC	TCT CTC AAT TCC CTC ATC ATC AAG CTG ATC CAG			60
50		60		360
Ala Asp Ile Leu Lys Asn Lys Leu Ser	Lys Gln Met Val Asp Val Lys Glu Asn Tyr Gln			360
GC GAT ATC TTG AAA AAC AAG CTT TCC	AAA CAG ATG GTG GAT GTT AAG AGA AAT TAC CAG			80
70		80		420
Ser Thr Leu Pro Lys Ala Glu Ala Pro	Arg Glu Pro Glu Gln Gly Glu Ala Thr Arg Ser			420
AGC ACC CTG CCC AAA GCA GAG GCA CCC	AGG GAA CCA GAG CAG GGA GAG GCC ACC AGG TCA			100
90		100		480
Glu Phe Gln Pro Met Ile Ala Thr Asp	Thr Glu Leu Leu Arg Gln Gln Arg Arg Tyr Asn			480
GAG TTC CAG CCA ATG ATT GCA ACG GAC	ACA GAG CTA CTA CGG CAA CAG AGA CGC TAC AAT			120
110		120		540
Ser Pro Arg Val Leu Leu Ser Asp Ser	Thr Pro Leu Glu Gln Pro Pro Leu Tyr Leu Met			540
TCG CCC CGG GTC CTG CTG AGT GAC AGC	ACC CCT TTG GAG CCC CCT CCC TTA TAC CTA ATG			140
130		140		600
Glu Asp Tyr Val Gly Asn Pro Val Val	Ala Asn Arg Thr Ser Pro Arg Arg Lys Arg Tyr			600
GAG GAT ATC TTG GGC AAC CCG GTG GTA	GCC AAT AGA ACC TCA CCA CCG AGG AAA CGC TAT			160
150		160		660
Ala Glu His Lys Ser His Arg Gly Glu Tyr	Ser Val Cys Asp Ser Glu Ser Leu Trp Val			660
GCA GAA CAT AAG AGT CAC CGA GGA GAG	TAC TCA GTG TGT GAC AGT GAG AGC CTG TGG GTG			180
170		180		720
Thr Asp Lys Ser Ser Ala Ile Asp Ile Arg	Gly His Gln Val Thr Val Leu Gly Glu Ile			720
ACC GAC AAG TCC TCA GCC ATT GAC ATT	CGG GGA CAC CAG GTC ACA GTG CTG GGG GAG ATC			200
190		200		780
Lys Thr GGT Asn Ser Pro Val Lys Gln Tyr	Phe Tyr Glu Thr Arg Cys Lys Glu Ala Arg			780
AAA ACC GGT AAC TCT CCT GTG AAA CAA	TAT TTT TAT GAA ACG AGA TGT AAA GAA GCC AGG			220
210		220		840
Pro Val Lys Asn Gly Cys Arg Gly Ile Asp	Asp Lys His Trp Asn Ser Gln Cys Lys Thr			840
CCG GTC AAA AAC GGT TGC AGG GGG ATT	GAT GAC AAA CAC TGG AAC TCT CAG TGC AAA ACT			240
230		240		900
Ser Gln Thr Tyr Val Arg Ala Leu Thr Ser	Glu Asn Asn Lys Leu Val Gly Trp Arg Trp			900
TCG CAA ACC TAT GTC CGA GCA CTG ACT	TCA GAA AAC AAC AAA CTC GTA GGC TGG CGC TGG			250
250		250		960
Ile Arg Ile Asp Thr Ser Cys Val Cys Ala	Leu Ser Arg Lys Ile Gly Arg Thr End			960
ATA CGA ATA GAC ACT TCC TGT GTG TGT	GCC TTG TCG AGA AAA ATT GGA AGA ACA TGA ATT			1030
1030		1030		1100
GGCATCTGTC CCCACATATA AATTATTACT	TTAAATTATA TGATATGCAT GTAGCATATA AATGTTTATA			1100
TTGTTTTTAT ATATTATAAG TTGACCTTTA	TTTATTAAAC TTCAGCAACC CTTACAGTAT ATAAGCTTTT			1170
TTCTCAATAA AATTCGTGTG CTGCTGCCCC	CTCAGGCCCT TCCATCTGT TAACCTTGTG TTGTGATTGG			1240
GCTCTGGGGA ACCCTCTCTG AAAACCTGTG	TACACACGTA TTGGGCATTC AGTATTGTCA AGGCCATGAC			1284
TTGTTGTTCA GTAAACTTTT TTAATAATCGG	ATGAGTCAGA GTTG			1284

and a 4,000-bp *Pst*I fragment, were subcloned into a Bluescript SK⁺ vector. The full *Hind*III sequence, 3'- and 5'-extended using the *Pst*I fragment, is shown.

from a variety of mouse tissues and analysed it with an NT-3-specific probe (Fig. 3). A single band of about 1.4 kilobases (kb) was found in all tissues examined (Fig. 3a). However, the level of expression of this mRNA varied considerably. In brain (Fig. 3b), there were marked regional differences in expression, with the highest mRNA levels occurring in cerebellum and hippocampus (Fig. 3b). These results demonstrate that the distribution of NT-3 mRNA is very different from that of BDNF and NGF mRNA in adult mouse tissue. Indeed, BDNF mRNA is found predominantly in the brain⁵, and NGF mRNA is hardly detectable in tissues such as liver or skeletal muscle¹⁰. We note that the hippocampus, which expresses NGF mRNA¹¹, also expresses BDNF and NT-3 mRNA, and in much larger amounts (Fig. 3b, and M. M. Hofer *et al.*, unpublished results).

V1		V2		
YAEKSHRGEYSVCDSESLWVT**DKSSAIDIRHQVTVLGEIKTGNSPVKQFYETRC				NT-3
SSTHPVFHMGESVCDVSVMV**GDKTTATDKEVTVLAENVINNSVFRQYFFETKC				NGF
HSDPARRGELSVCDSEIENVTADDKTAVDMSGGTVTVLEKVPVSKGLKQFYETKC				BDNF
V3		V4		
KEARPVKNGCRGIDDKHNSQCKTSQTYVRALTSENKLVGMWRIRIDTSCVCLSRKIGRT				NT-3
RASNPVSGCRGIDDKHNSYCTTTHTFVKALTTDEKQ*AAARFIRIDTACVCLSRKATRRG				NGF
NPMGYTKECRGIDDKHNSQCKRTTQSYVRALTMDSKRRIGRIRIDTSCVCLTIKRRG				BDNF

FIG. 2 Amino-acid (single-letter code) sequence comparison of mouse NT-3, NGF¹⁹ and BDNF⁶. The mouse BDNF sequence is 100% identical with that of pig BDNF (A.H. *et al.*, unpublished results). Bold letters indicate identical amino acids found at the same positions in all three proteins; arrows indicate the six cysteine residues. Asterisks indicate gaps introduced to optimize matching. V1–V4, the four variable domains of more than three contiguous amino acids.

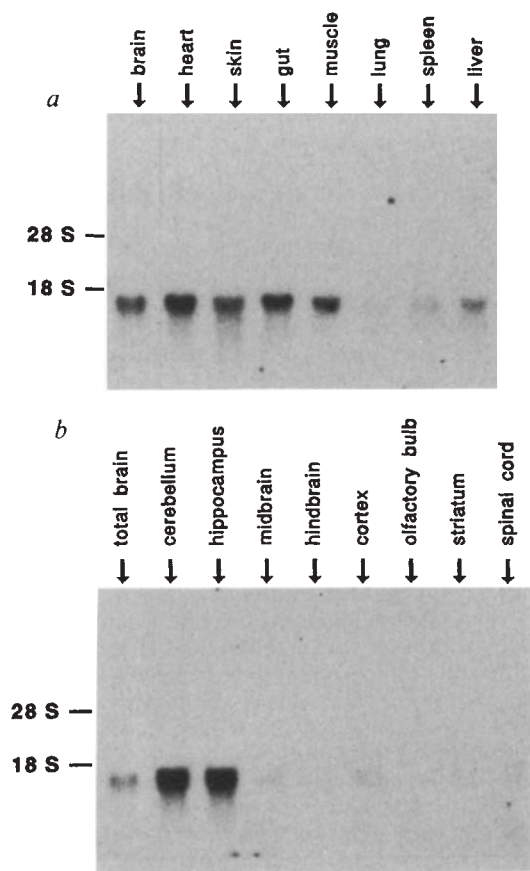


FIG. 3 Tissue distribution of NT-3 mRNA. Total RNA (20 μ g) was applied to each lane and hybridized with a 32 P-labelled double-stranded DNA probe. A single band corresponding to ~ 1.4 kb can be seen in all tissues, the weakest signal being observed in lung, and the strongest, in heart (a). Skeletal muscle was from the thighs. In brain (b), the strongest signal was in hippocampus and cerebellum. Ribosomal RNA size markers are shown on the left.

METHODS. Total RNA was extracted from adult female tissues¹⁸, and electrophoresed on 1.3% formaldehyde-containing agarose gels²⁰. The RNA was transferred to nylon membranes (Hybond-N, Amersham) and hybridized overnight at 42 °C in 1 ml sodium phosphate (200 mM, pH 7.2), containing 40% formamide, 5 $\times$ Denhardt's solution and 200 μ g ml⁻¹ denatured salmon-sperm DNA. A 32 P-labelled random-primed double-stranded probe²¹ corresponding to nucleotides 319–1,093 was used (Fig. 1). The specific activity of the probe was 1.3×10^8 c.p.m. μ g⁻¹, and 10^7 c.p.m. ml⁻¹ were added to the hybridization buffer. Washing was for 60 min at 60 °C in 0.1 $\times$ SSC containing 0.5% SDS. Filters were exposed for 5 days at -70 °C with intensifying screens.

To test whether NT-3 is a biologically active and secreted protein, the entire protein coding sequence was cloned into an expression vector (plasmid pCMV) and introduced into monkey kidney COS cells (Fig. 4). Because NT-3 mRNA is found in peripheral tissues, we cultured a variety of embryonic chick neurons that project to these tissues, and require trophic factors for survival. Motorneurons, purified from the spinal cord of 6-day-old embryos (E6)¹², ciliary neurons (E8) and sympathetic neurons (E11), did not survive in the presence of NT-3 (data not shown). However, NT-3 supported the survival of about 30% of neurons isolated from the nodose ganglion (Fig. 4). Furthermore, this effect was additive to that of BDNF, which acts on a subpopulation of the nodose neurons¹³, and the combination of NT-3 and BDNF resulted in the rescue of most (90%) neurons (Fig. 4). It is important to note that NT-3 mRNA is found in the visceral targets of this ganglion, including heart, gut, liver and lung (Fig. 3a). Whereas only small numbers of NT-3-responsive neurons were found in E8 dorsal root and trigeminal ganglia, essentially all neurons (94%, s.d. ± 12 n = 3) isolated from the E10 trigeminal mesencephalic nucleus were

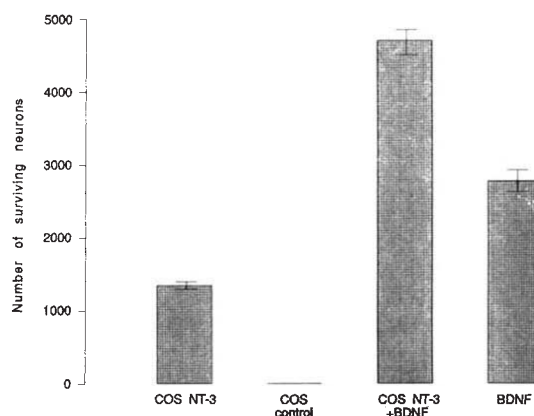


FIG. 4 Survival of sensory neurons isolated from nodose ganglia obtained from 8-day-old embryonic chick. Cells (5,000) were plated on a polyornithine-laminin substrate, and the surviving neurons counted after 24 h. Results are means of triplicate determinations $\pm$ s.d. The concentration (1 ng ml⁻¹) of BDNF used was three times the minimal concentration required to obtain maximal survival. The NT-3 COS medium was used at a solution of 1:50. No neurons survived without addition of NT-3 or BDNF, or with conditioned medium used at a 1:50 dilution with nontransfected COS cells, or cells transfected with control DNA (see below).

METHODS. Oligonucleotide primers were synthesized corresponding to the first 19 nucleotides (plus an *Eco*RI site) and last 19 nucleotides (plus a *Bam*HI site) of the open reading frame shown in Fig. 1. After PCR amplification using a *Pst*II genomic fragment as template (Fig. 1), the resulting product was cloned into the *Eco*RI-*Bam*HI site of the expression vector plasmid pCMV²². The nucleotide sequence of the NT-3 insert in plasmid pCMV was determined. COS-1 cells were transfected using the calcium phosphate method²³ and the medium collected as previously described⁶. The control media used were obtained after treatment of the COS-1 cells with calcium phosphate, or with a pCMV-NT-3 construct in which the stop codon of NT-3 was deleted. Both media had no biological activity at 1:50 dilution. Nodose ganglia were dissociated, and the neurons cultured in 24-well plates¹³. BDNF was purified from pig brain²⁴. Neurons from E10 trigeminal mesencephalic nuclei were isolated as described¹⁵, with minor modifications. 600 neurons (as determined 3 h after plating) were plated on a polyornithine-laminin substrate in 24-well plates and counted after 24 h (see text).

found to survive in the presence of NT-3. These proprioceptive neurons innervate skeletal muscle, a tissue found to express NT-3 mRNA (Fig. 3a). It therefore seems that NT-3 represents the biological activity present in several peripheral tissues, including liver¹⁴ and skeletal muscle¹⁵, and known to support the survival of visceral and proprioceptive sensory neurons that do not respond to NGF (see ref. 2 for a review). □

Received 24 January; accepted 14 February 1990.

1. Purves, D. in *Body and Brain. A Trophic Theory of Neural Connections* (Harvard University Press, Cambridge, 1988).
2. Davies, A. M. *Development* **101**, 185–208 (1987).
3. Barde, Y.-A. *Neuron* **2**, 1525–1534 (1989).
4. Cohen, S. *Proc. natn. Acad. Sci. U.S.A.* **46**, 302–311 (1960).
5. Barde, Y.-A., Edgar, D. & Thoenen, H. *EMBO J.* **1**, 549–553 (1982).
6. Leibrock, J. et al. *Nature* **341**, 149–152 (1989).
7. Saiki, R. K. et al. *Science* **230**, 1350–1354 (1985).
8. Edwards, R. H., Selby, M. J., Garcia, P. D. & Rutter, W. J. *J. biol. Chem.* **263**, 6810–6815.
9. Angeletti, R. H., Hermodson, M. A. & Bradshaw, R. A. *Biochemistry* **12**, 100–115 (1973).
10. Heumann, R., Korsching, S., Scott, J. & Thoenen, H. *EMBO J.* **3**, 3183–3189 (1984).
11. Korsching, S. et al. *EMBO J.* **4**, 1389–1393 (1985).
12. Dohmann, U. et al. *Dev. Biol.* **118**, 209–221 (1986).
13. Lindsay, R. M., Thoenen, H. & Barde, Y.-A. *Dev. Biol.* **112**, 319–328 (1985).
14. Lindsay, R. M. & Tarbit, J. *Neurosci. Lett.* **12**, 195–200 (1979).
15. Davies, A. M. *Dev. Biol.* **115**, 56–67 (1986).
16. Innis, M. A., Myambo, K. B., Gelfand, D. H. & Brow, M. A. D. *Proc. natn. Acad. Sci. U.S.A.* **85**, 9436–9440 (1988).
17. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
18. Okayama, H. et al. *Meth. Enzym.* **154**, 3–28 (1987).
19. Scott, J. et al. *Nature* **302**, 538–540 (1983).
20. Lehrach, A., Diamond, D., Worney, J. & Boedtker, H. *Biochemistry* **16**, 4743–4751 (1977).
21. Feinberg, A. P. & Vogelstein, B. *Analyt. Biochem.* **137**, 266–267 (1979).
22. Andersson, S. et al. *J. biol. Chem.* **264**, 8222–8229 (1989).
23. Chen, C. & Okayama, H. *Molec. cell. Biol.* **7**, 2745–2752 (1987).
24. Hofer, M. M. & Barde, Y.-A. *Nature* **331**, 261–262 (1988).

ACKNOWLEDGEMENTS. We thank Hans Thoenen for comments on the manuscript, Yoshi Arakawa for performing the motorneuron experiments, and Frau Claudia Cap for technical assistance. This work was supported by the Bundesministerium für Forschung und Technologie. Y.A.B. is supported by the Hermann-und-Lilly Schilling Stiftung.

Retinofugal fibres change conduction velocity and diameter between the optic nerve and tract in ferrets

Gary E. Baker & Michael P. Stryker*

Department of Human Anatomy, University of Oxford, South Parks Road, Oxford OX1 3QX, UK

IN earlier studies of central nervous fibre tracts, it was tacitly assumed that individual axons are relatively uniform along their length. In the retinofugal pathway in particular, axon diameter, myelin thickness and correlated conduction properties^{1,2} have been treated as constant throughout the optic nerve, chiasm and tract. We report here that the conduction velocities of fibres contributing to the early components of the compound action potential are significantly greater in the optic tract than in the optic nerve of ferrets, and also that the diameters of the largest retinofugal fibres increase from nerve to tract³. This observation raises significant questions about the developmental mechanisms in the central nervous system that relate the axons, their diameters, and the glia with which they are myelinated. In addition, it indicates that studies that have relied on the constancy of conduction velocity along the retinofugal course⁴ may require reappraisal.

Five normally pigmented adult ferrets were studied using the arrangement of electrodes shown in Fig. 1*a*. Records of the compound action potential (Fig. 1*b–g*) were obtained with electrodes positioned: intraretinally near the optic disk; around the optic nerve about 1 mm behind the eye; in the optic nerve about 3 mm distal to its point of fusion with the contralateral nerve; in the optic chiasm contralateral to the midline; and at the contralateral optic tract close to the lateral geniculate nucleus (Fig. 1*a*). Each record shows two major conduction latency modes, one (t_1) with a short latency negative peak, and a second (t_2) with a slower peak⁵. This form of compound action potential is similar to that previously found in studies of the cat retinofugal pathways^{6,7}. Each conduction latency mode in the compound action potential is believed, on the basis of studies of the peripheral nervous system⁸ and various investigations of the retinofugal pathways^{9–11}, to reflect the presence of a subpopulation of fibres with similar conduction velocity.

The t_1 and t_2 conduction velocities were determined for the optic nerve and optic tract from measurements of conduction times and distances. Conduction times for the nerve and tract were calculated by two methods. First, from recordings made with the intraretinal electrode (RET) near the optic disk, we took the tract conduction time to be the difference between the latencies of the peak t_1 and t_2 negativities elicited from optic tract (OT) and optic chiasm (OX) stimulation. Similarly, we took the nerve conduction time to be the difference between the latencies of the potentials elicited by intracranial optic nerve (ONc) and intraorbital optic nerve (ONo) stimulation. Figure 1*h* shows that these measurements depend only negligibly on stimulus intensity over the range used. The measured distances were divided by these conduction times to yield a single estimate of conduction velocities. These measures of latency incorporate a period of delay, the utilization time, between stimulus onset and spike initiation in the stimulated fibres. This utilization time has been described as being negligibly small for optic tract stimulation in the cat¹² and ~100 μ s in the rat optic nerve¹³. If the utilization time is equal at all of the stimulating electrodes, it would have no effect on the conduction velocities measured by this method, as the same delay would be subtracted from

each latency. This measure of conduction velocity is largely independent of which points on the t_1 and t_2 waveforms are chosen for the latency measurement, as similar waveforms are measured in the various stimulus conditions. Figure 2*a, b* shows that the t_1 velocities in the optic tract were consistently much greater than those in the nerve. The difference between nerve and tract velocities was smaller and more variable for the t_2 responses, but the measured tract velocities were greater than or equal to those in the nerve in every case.

If the utilization times do differ greatly for the different stimulating electrode placements, the previous method may be invalid¹⁴. We therefore measured conduction times by a second method, recording the peak t_1 and t_2 negativities of antidromic responses at ONo and orthodromic responses at OT after identical OX stimulation (Fig. 1*f, g*). Using this strategy, a single utilization time is common to latencies measured in both nerve and tract. The optic tract velocities measured by this method were consistently greater than those in the nerve, even when the utilization time is assumed to be zero. The difference between the mean values of tract and nerve velocities was 36% for the t_1 mode and 20% for the t_2 mode. These differences would, of course, increase if a realistic value was assumed for the utilization time. Therefore, although the increase from nerve to tract was smaller when measured by this second method, indicating that the utilization time may not have been constant along the retinofugal pathway as assumed for the previous method, the differences between nerve and tract illustrated above are not artefacts of unequal utilization times.

Figure 3*a–e* shows that in each of the animals studied physiologically, there were many more axons with a diameter >5 μ m in the optic tract than in the optic nerve. In the ferret optic tract large-calibre axons are known to be of retinal origin³. If we assume that the largest axons in the nerve increase their diameters to become the largest axons in the tract, we can estimate the magnitude of this increase graphically, as in Fig. 3*f*, by shifting the nerve axon population to the right until it is superimposed on the equivalent number of largest axons in the tract. For an axon of 5 μ m diameter in the nerve, the average increase in diameter estimated by this method was 16.9% (range, 5–40%, $n = 5$). The data of Figs 2 and 3 show, therefore, that the largest and fastest optic tract axons are consistently larger and faster than those in the nerve, with no overlap between the two samples. These data do not determine the velocity–diameter relationship, however, because the anatomical data were obtained from the extreme upper tail of the axon diameter distribution, whereas the physiological measurements probably reflect the majority of the many thousands of axons in each axon diameter class.

In the retinofugal pathway of the cat, each of the peaks of the compound action potential has been associated with a physiologically and anatomically distinct cell type, the fastest (t_1), with large diameter, called Y- or α -cell fibres and the slower (t_2), with predominantly intermediate diameter, called X- or β -cell fibres^{15–17}. We have not studied the slowest (t_3) latency group, associated with W-cell fibres^{18,19}. It is therefore reasonable to infer from our data that the conduction velocities of both Y- and X-cells increase along their central course. Consequently, calculations of conduction velocity from measurement of conduction time along only a segment of the retinofugal pathway are likely to be inaccurate. For example, to infer the total extraretinal conduction time from measures of conduction velocity in the optic tract may underestimate the total conduction time⁴.

In the developing ferret optic nerve, glia have an interfascicular arrangement distally and a radial organization more centrally²⁰, and different distributions of glial precursor cells have been identified between the eye and chiasm in the developing optic nerve of rats²¹. The fascicular arrangement of fibres in the adult optic nerve changes abruptly to a non-fascicular organization just rostral to the chiasm, with a partial segregation

* Present address: Department of Physiology and Neuroscience Graduate Program, University of California, San Francisco, California 94143-0444, USA.

of fibre diameter types similar to that in the optic tract²². Each retinal-ganglion-cell axon thus passes through three different glial environments. It has long been recognized that the conduction velocity in the unmyelinated, intraretinal segment is much lower than that in the more central segment¹⁹. The results presented here show that conduction velocity and axon diameter also differ between the nerve and tract environments. We suggest that the change in axon diameter may be a result of the change

in glial type as the axons pass from one region of the central pathway to the other, but the precise region in which the change in diameter occurs remains to be defined. It will, therefore, be interesting to determine whether the change described here is gradual, extending over a considerable distance, or local, occurring at the site where the glial environment of the axons changes.

The stable relationship between axon diameter and myelin thickness over the course of regeneration and development has

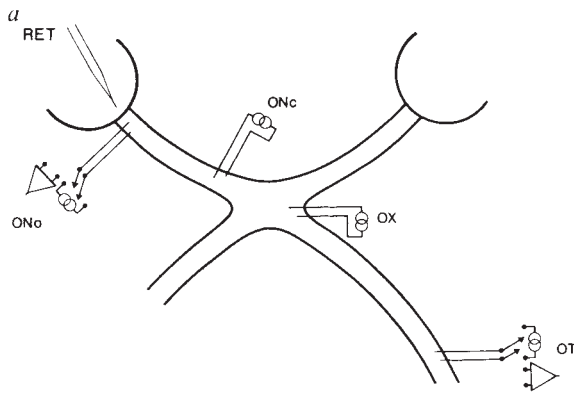
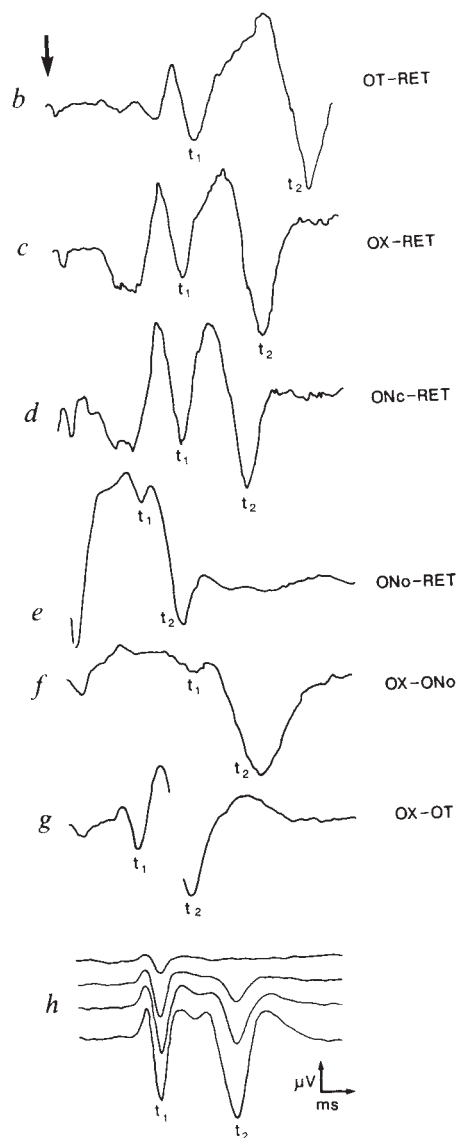


FIG. 1 *a*, Schematic drawing to illustrate the arrangement of electrodes. *b–e*, Antidromic potentials recorded at the optic disk (RET) following stimulation of: *b*, optic tract (OT); *c*, optic chiasm (OX); *d*, intracranial optic nerve (ONc); *e*, intraorbital optic nerve (ONo). *f*, Antidromic potential recorded at ONo following stimulation at OX. *g*, Orthodromic potential recorded at OT following stimulation at OX. Conduction distances (in this animal): OT – ONo, 29.5 mm; OX – ONo, 17.9 mm; ONc – ONo, 13.4 mm. *h*, Series of antidromic potentials recorded at RET following stimulation of increasing intensity at OT. Note that the conduction latencies for t_1 and t_2 remain nearly constant throughout the series. Calibration: *b, c, d*, 20 μ V, 0.5 ms; *e*, 50 μ V, 0.5 ms; *f*, 20 μ V, 0.2 ms; *g*, 50 μ V, 0.2 ms; *h*, 200 μ V, 0.5 ms. Stimulus onset is indicated by arrow.

METHODS. Anaesthesia was induced with xylazine (Rompun; 1 mg kg⁻¹) and ketamine (Vetalar; 15 mg kg⁻¹) intramuscularly and maintained by intravenous sodium pentobarbitone (Sagatal; 10 mg kg⁻¹). During recording, animals were placed in a stereotaxic frame and rectal temperature was regulated at $38.0 \pm 0.5^\circ\text{C}$. After the tissues behind the globe were dissected, a bipolar electrode manufactured from two pieces of Teflon-coated stainless steel wire ($7 \times 125 \mu\text{m}$ outer diameter) stripped of the insulation at the tips, was hooked onto the intraorbital optic nerve. A concentric bipolar iridium electrode (Rhodes) was advanced through the vitreous body onto the region of the optic disk through a scleral puncture. Craniotomies were made and, after reflection of the dura, bipolar electrodes were positioned at the optic chiasm and tract using coordinates previously computed for the ferret³⁰. The reference electrode for recording was placed in the exposed tissue of the scalp. The positions of intracranial stimulating electrodes were finally adjusted to evoke the lowest threshold responses at the optic disk. Compound action potentials were recorded (a.c. coupled, 0.1 Hz–10 kHz) using a high impedance head stage (Neurolog NL100) and displayed on a storage oscilloscope (Tek 5113) from which they were photographed (Polaroid C5). Conduction latencies from stimulus onset to the negative peak of each latency mode were measured directly from the photographic records using a digital caliper. Responses were recorded at each electrode following constant current stimuli (50 or 100 μsec duration; variable intensity) delivered to the other electrode sites. Electrode configuration: concentric bipolar electrodes were used for RET, ONc, OX and in some experiments for OT. ONo was a colinear bipolar electrode inserted normally to the course of the optic nerve, with $<250 \mu\text{m}$ tip separation parallel to the course of the optic nerve. When the OT electrode was not concentric bipolar, it was a parallel bipolar electrode (with the tips separated by 2 mm in the coronal plane) lowered on a vertical axis to stimulate optic tract axons rising nearly vertically. All other electrodes (OX, ONc, ONo) were advanced along a course normal to the axons stimulated. Electrode position and effects of stimulus polarity: in every case, electrodes were sited at the low threshold point with the cathode actually in or on the optic nerve or tract (confirmed histologically). Cathodal stimulation was always effective at lower (usually very much lower) intensities than anodal stimulation. In cases in



which the OT electrode was parallel bipolar instead of concentric bipolar, unipolar stimulation (using a distant extracranial reference) was found to yield identical latencies over the range of stimulus intensities used for measurement of responses. Effects of stimulus intensity: at every site, a series of stimulus intensities was used, ranging from subthreshold through gradually increasing steps, to suprathreshold for both the t_1 and t_2 responses. The records measured were those of the lowest intensity at which a clear response was observed for the component measured. Thus, in many cases, the t_1 response rose to maximum (and in some cases shortened in latency) before the t_2 response had attained moderate size; in such cases we measured the t_1 and t_2 latencies from different records, each at an intensity at which it had attained less than its half-maximal size. We usually observed very little latency shift during manifold increases in stimulus intensity. Stimulus spread: none of the range of stimulus intensities used in this study was capable of activating optic axons at shorter latencies from OX than from ONc, indicating that stimulus spread equal to or greater than the distance separating them did not take place from those electrodes.

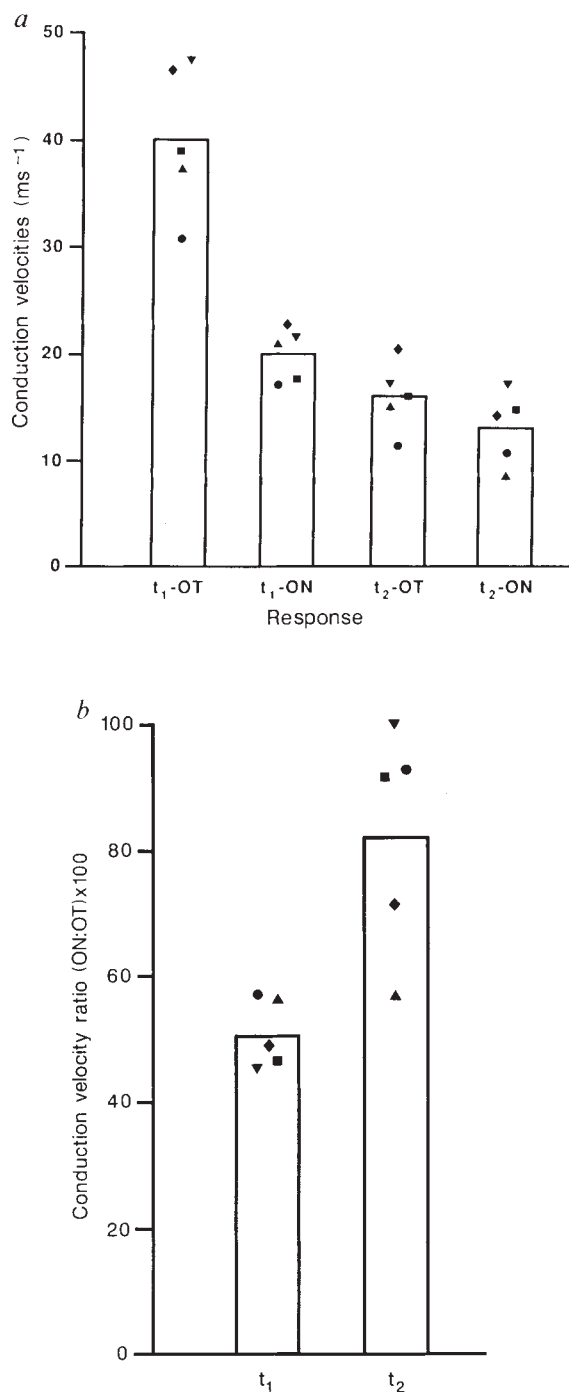


FIG. 2 *a*, Conduction velocities for t_1 and t_2 responses in optic nerve (ON) and tract (OT). *b*, Conduction velocities in optic nerve as a percentage of velocity in the optic tract for each animal. Bars show the means; individual animals are coded respectively by a triangle, inverted triangle, square, diamond and circle.

METHODS. At the end of recording, electrolytic lesions were made at the tips of the centrally sited electrodes, and animals were perfused with a solution of 3% paraformaldehyde and 1% glutaraldehyde in 0.1 M phosphate-buffered saline. Segments of optic nerve and tract contralateral to those from which recordings had been made were removed and retained for subsequent fibre analysis. Conduction distances were measured along the course of the optic nerve and tract from the back of the globe to visible electrode sites and landmarks (fusion of the optic nerves, stereotactically directed knife cuts) using a length of fine silk thread. For measurements of distances between deep sites, blocks of tissue were embedded in gelatin-albumin, sectioned on a Vibratome at 100 μm and stained with cresyl violet. The course of the optic tract was reconstructed using an X-Y microscope stage encoder (Boekeler) and the distance separating lesion sites calculated.

been interpreted as indicating that growing axons must inform the associated glial cells of just how many turns of myelin to produce^{23,24}. The observations reported here indicate that, as for dorsal root ganglia *in vitro*²⁵ and *in vivo*²⁶, this interaction seems to be a two-way affair, with the glial environment influencing the final axon diameter as well.

Changes of conduction velocity and diameter have been reported between peripheral nerve roots and the spinal cord^{27,28} but the identity and origin of the myelinating cells, Schwann cells peripherally and oligodendrocytes centrally, changes completely at the central-peripheral boundary²⁹. The change observed here takes place entirely within the central nervous system along a fibre pathway that has generally been regarded

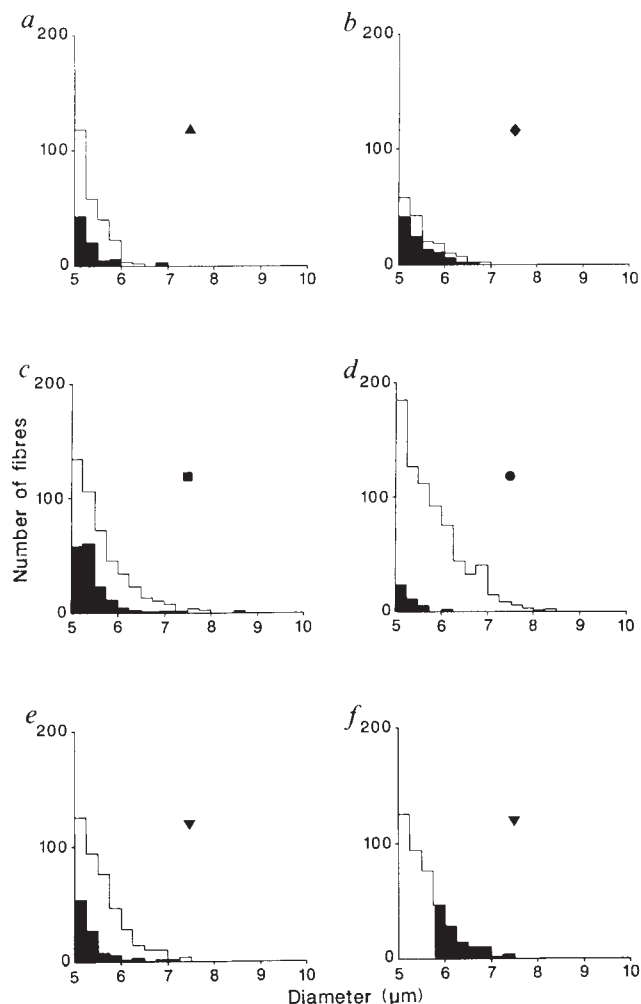


FIG. 3 *a-e*, Diameter distribution of all axons with an internal diameter of $>5 \mu\text{m}$ in the optic nerve (shaded) and tract of each of the animals that were studied physiologically. *f*, Optic nerve axon diameter distribution in the animal shown in *e*, shifted rightward by 0.81 μm to match the distribution of the largest axons in the optic tract.

METHODS. Optic nerve and tract segments were sectioned transverse to their course at 200 μm on a Vibratome, post-fixed in 1% osmium tetroxide, dehydrated through graded ethanols and embedded in Epon. Semi-thin sections (1 μm) were stained with 1% *p*-phenylenediamine. The internal circumferences of all axons of internal diameter $\geq 4 \mu\text{m}$ (as judged by comparison with standard circles) in each nerve and tract were drawn from semi-thin sections at magnification $\times 1,500$ using a drawing tube attached to a microscope. The circumference of each profile drawn was then measured using an image analysis system (Kontron IPS), which also calculated the diameters and constructed the histograms. The data presented here are restricted to axons with an internal diameter $>5 \mu\text{m}$, to eliminate the possibility of incomplete sampling.

as representative of all central nervous tracts. We must now ask whether similar changes in conduction velocity and diameter take place in other central pathways. □

Received 29 November 1989; accepted 16 January 1990.

1. Hursh, J. P. *Am. J. Physiol.* **127**, 131–139 (1939).
2. Waxman, S. G. & Swadlow, H. A. *Prog. Neurobiol.* **8**, 297–324 (1977).
3. Baker, G. E. & Guillery, R. W. *Soc. Neurosci. Abstr.* **14**, 992 (1988).
4. Stanford, L. R. *Science* **238**, 358–360 (1987).
5. Bishop, P. O. & Mcleod, J. G. *J. Neurophysiol.* **17**, 387–414 (1954).
6. Bishop, P. O., Jeremy, D. & Lance, J. W. *J. Physiol. Lond.* **121**, 415–432 (1953).
7. Chang, H. T. *J. Neurophysiol.* **19**, 224–231 (1956).
8. Erlanger, J. & Gasser, H. S. *Electrical Signs of Nervous Activity* 2nd edn (University Pennsylvania Press, Philadelphia, 1968).
9. Cleland, B. G., Dubin, M. W. & Levick, W. R. *J. Physiol. Lond.* **217**, 473–496 (1971).
10. Stone, J. & Fukuda, Y. *J. Neurophysiol.* **37**, 722–748 (1974).
11. Stone, J. & Hoffman, K.-P. *Brain Res.* **43**, 610–616 (1972).
12. Stone, J. & Freeman, R. B. *Exp. Brain Res.* **13**, 489–497 (1971).
13. Sumitomo, I., Ide, K., Iwama, K. & Arikuni, T. *Exp. Neurol.* **25**, 378–392 (1969).
14. Ranck, J. B. *Brain Res.* **98**, 417–440 (1975).
15. Cleland, B. G., Levick, W. R. & Wässle, H. *J. Physiol. Lond.* **248**, 151–171 (1975).
16. Rowe, M. H. & Stone, J. *Exp. Brain Res.* **25**, 339–357 (1976).
17. Rowe, M. H. & Stone, J. *Brain Behav. Evol.* **14**, 185–216 (1977).
18. Bishop, G. H., Clare, M. H. & Landau, W. M. *Exp. Neurol.* **24**, 386–399 (1969).
19. Stone, J. & Fukuda, Y. *J. Neurophysiol.* **37**, 722–748 (1974).
20. Guillery, R. W. & Walsh, C. *J. comp. Neurol.* **265**, 203–217 (1987).
21. Raff, M. C. *Science* **243**, 1450–1455 (1989).
22. Baker, G. E. *Eur. J. Neurosci.* **2**, 24–33 (1990).
23. Simpson, S. A. & Young, J. Z. *J. Anat.* **79**, 48–65 (1945).
24. Friede, R. L. *J. comp. Neurol.* **144**, 233–252 (1972).
25. Windebank, A. J., Wood, P., Bunge, R. P. & Dyck, P. J. *J. Neurosci.* **5**, 1563–1569 (1985).
26. Pannese, E., Ledda, M. & Matsuda, S. *J. Neurocytol.* **17**, 693–700 (1988).
27. Fraher, J. P. *J. Anat.* **126**, 509–533 (1978).
28. Traub, R. J. & Mendell, L. M. *J. Neurophysiol.* **59**, 41–55 (1988).
29. Maxwell, D. S., Kruger, L. & Pineda, A. *Anat. Rec.* **164**, 113–126 (1969).
30. Zahs, K. & Stryker, M. P. *J. comp. Neurol.* **241**, 210–224 (1985).

ACKNOWLEDGEMENTS. We thank R. W. Guillery for reading the manuscript, Terry Richards for artwork and Ann Stanmore for assistance with the image analysis system. This work was supported by the MRC. M.P.S. was supported by the University of California, the McKnight Foundation, and the NIH.

Porcine diacylglycerol kinase sequence has zinc finger and E-F hand motifs

Fumio Sakane*, Keiko Yamada*, Hideo Kanoh*, Chieko Yokoyama† & Tadashi Tanabe†

*Department of Biochemistry, Sapporo Medical College, West-17, South-1, Sapporo 060, Japan

†Department of Pharmacology, National Cardiovascular Center Research Institute, Fujishiro-dai, Suita, Osaka 565, Japan

CELL stimulation causes diacylglycerol kinase (DGK) to convert the second messenger diacylglycerol into phosphatidate, thus initiating the resynthesis of phosphatidylinositols and attenuating protein kinase C activity¹. Of the DGK isoforms so far reported^{2–4}, only porcine DGK from lymphocytes⁵ has been characterized in detail^{3,5–7}. Here we report the isolation and sequencing of complementary DNA clones that together cover the entire region encoding porcine DGK (relative molecular mass 80,000 (80K)). The deduced primary structure of this DGK contains the putative ATP-binding sites, two cysteine-rich zinc finger-like sequences similar to those found in protein kinase C⁸, and two E-F hand motifs, typical of Ca²⁺-binding proteins like calmodulin⁹. Indeed, we find that the activity of this DGK isoform is enhanced by micromolar concentrations of Ca²⁺ in the presence of deoxycholate or sphingosine. These properties of 80K DGK indicate that its action is probably linked with both of the second messengers diacylglycerol¹⁰ and inositol 1,4,5-trisphosphate¹¹.

We screened porcine thymus and spleen cDNA libraries with synthetic oligodeoxynucleotides (whose sequence was based on a partial primary structure of purified DGK) and a cloned cDNA fragment as probes. Six positive cDNA fragments were then sequenced as shown in Fig. 1a: the composite sequence (Fig.

1b) contains an open reading frame encoding 734 amino acids (including the initiating methionine) whose sequence includes six lysylendopeptidase peptides of purified DGK (underlined in Fig. 1b). The calculated relative molecular mass (M_r) of the encoded protein (82,605) agrees with determinations by SDS-PAGE for the purified enzyme (78–83K)^{3,6}. The nucleotide sequence around the initiation codon fulfills Kozak's criteria for eukaryotic initiation sites¹². In the 5'-untranslated region of the cDNA, there are in-frame stop codons at positions –96, –186, –258 and –279; a polyadenylation signal (AATAAA) appears at nucleotides 2,453–2,458 and there is a poly(A) tract at the 3'-terminus. Northern hybridization of total RNA from several tissues shows that the size of the 80K DGK messenger RNA is ~2.9 kilobases (kb) (Fig. 2a), which is slightly higher than the composite length of the cDNAs (2,772 bases). The mRNA is most abundant in thymus and is present in brain and other tissues to a lesser extent and is undetectable in liver (Fig. 2a), a result confirmed by immunoblotting^{2,5}.

Using a phagemid pBluescript, we then constructed a fusion gene by ligating a *lacZ* gene containing the promoter region to the λ DGK1 insert, which lacks the segment encoding the N-terminal 13 amino acids (compare Fig. 1a). This fusion gene, inserted into the phagemid, was used to transfect *Escherichia coli*, and expression of the gene induced with isopropyl- β -D-thiogalactoside. A new protein band was detected in an extract of the transfected cells by western blotting with anti-80K DGK antibodies⁷ (Fig. 2b); this band was not found in the cells transfected with the vector alone. The apparent M_r of this protein (83K), which is slightly higher than the 80K of DGK purified from porcine thymus³, corresponds to the M_r expected for the fusion protein (85.4K). In view of the immunospecificity of the antibodies^{2,5,7}, our cDNA must encode the 80K DGK. Furthermore, DGK activity in the extract from fusion gene-transfected cells (0.9 nmol min⁻¹ mg⁻¹ protein) was significantly higher than from cells transfected with vector alone (0.2 nmol min⁻¹ mg⁻¹ protein).

The primary structure predicted for 80K DGK contains two cysteine-rich, zinc finger-like sequences at residues 218–252 and 281–318. Such sequences have been found in DNA-binding proteins involved in transcription regulation¹³ and also in the protein kinase C proteins (PKCs)⁸. Interestingly, the sequences detected in the DGK (Cys-X₂-Cys-X₁₂-Cys-X₂-Cys-X₇-Cys-X₆-Cys; and Cys-X₂-Cys-X₁₄-Cys-X₂-Cys-X₇-Cys-X₇-Cys, where X is any amino acid) are similar to those occurring in PKCs⁸. More than 30% of the first sequence is identical with the two cysteine-rich sequences of the α -subspecies of PKC, whereas the identity is ~20% between the second DGK sequence and those of PKC (Fig. 3a). The cysteine-rich sequences of PKCs are involved in phorbol ester binding¹⁴, but it is not known whether DGK can serve as a phorbol ester receptor. In a preliminary experiment, we have found that phorbol 12-myristate 13-acetate at 1.0 μ M inhibits DGK activity to a limited extent (10–15%) when assayed with 50 μ M diolein, which gives half-maximal activity⁶ for the 80K DGK, but we cannot say yet whether 80K DGK can also bind phorbol esters. Although it is not known if DGK binds DNA, we have observed dense staining of neuronal nuclei with anti-80K DGK antibodies⁷. The presence of zinc finger-like sequences in 80K DGK indicates that this enzyme and the PKCs¹⁵ could share similar regulatory mechanisms. Both enzymes are known to undergo intracellular translocation^{1,8,10}.

An unexpected feature of the DGK primary structure is that it contains two E-F hand motifs, which constitute Ca²⁺-binding sites in calmodulin⁹ and Ca²⁺-activated neutral protease¹⁶ for example. A search for the E-F hand structure in the DGK sequence using a test sequence consisting of 16 characteristic residues^{9,17} reveals two highly scored sequences at residues 113–141 (13 scores per 16 residues (13/16)) and 158–186(11/16). The secondary structure characteristic of the E-F hand motif, α -helix-loop- α -helix¹⁷, is also evident from Chou-Fasman

of E-F hand motifs indicates that it could be modulated by Ca^{2+} . As shown in Fig. 4, the activity is actually enhanced by 10^{-6} – 10^{-4} M Ca^{2+} in the presence, but not in the absence, of deoxycholate⁶ or sphingosine³; higher Ca^{2+} concentrations were inhibitory. This effect is not due to Ca^{2+} –ATP complex formation as it is independent of ATP concentration up to 3 mM. DGK

a

Genomic map showing the 2.1 kb BamHI-HincII-BamHI fragment and its subclones. The top part shows a linear map with restriction sites at 381, 1141, and 1317 bp. Below are diagrams of subclones pDGK 21, 23, 29, 33, 32, and λDGK1, with arrows indicating the direction of transcription or cloning.

b

-281
-240
-120

CTAGGTCCTCAGTTCCTACCAGTAAGTCCTAGGCCTCTG
TCTGTCTCCTTGCCTTACCAGCCTTACCAGTATCATGCTGTTCCAAGTGCTTTAACTCCAGAATAGATACCAGACGGGAGGACGCGAAGAGTCTCTTACCACGAGCTTTGGCGAGT
GCTTTCTACACACCTCCGCAAGACTAGCAGGCCACCCCTCGAAGCAGAGTCCCGGCACCGCTCCAGAGGCGCTACTACCAGCGTGTCTTTCTGTGTCATTCTTAAGAGAAAAAAG

1
1

ATGTCCTCAAGAGAGGGGGCTGTAAAGCCCACTGATTTTGCCACAGCTGCAAAAAATACATGAATACTCCACAAAAGGTCAAGTAGTGCTCTAAAGCTCTTTGAGGATGGTGAATGGCT
M S K E R G L I S P S D F A Q L Q K Y M E Y S T K K V S D V L K L F E D G E M A

121
41

GAATATCTCCAGGAGATGCCATTGGGTACGAGGGATTCAGCAACTCTGAAATCTATCTGGAAGTGGATAGTGTTCTCTAGTCAAGCTTAAGCGCTGTTTCAATCTCCAGACT
E Y L Q G G D A I G Y E G F Q Q F L K I Y L E V D S V P S H L S L A L F Q S S F Q T

241
81

AGTTACTGCTCAGAAGAGCTGTAAAAAGAGATGTGGTGTGCTCAGTGATGTCTCTGCTACTTCTCCCTTCTGAGGCGGCGCGCAGAGACAAGCTAGAGTTACCTTCAAGCTG
S Y C E K E A E T V K R D V V C L S D V S C Y F S L L E G G R P E D K L E F T F K L

361
121

TAGACACGACGAAATGGGATCTGGACAGCTCAGAAGCTGGACAGAAATCATCATACAGATGATGCGAATGGCTGAATAGCTGGATGGGTGCTGAGCTGAGCGCGGATCTCAG
Y D T D R N G I L D S S E V D R I I Q M M R M A E Y L D W D V S E L R P I L Q

481
161

GAGATGTAGAAAGAGATTGACTATGATGGAAGTGCTGTCTCCCTAGCTGAGTGCGCTCGGGCTGGGGCCACCACTGTGCCATTGCTTGTGTGCTGGGTCTGGAGTAGACCTTGAA
E M M K E I D Y D G S G S V S L A E W L R A G A T T V P L L G V G G M M T L K

601
201

GACAATGGGACGACATGTGGAGACCAAGAGCTTCCCGACCACTCATGCAAGCTGTGCGAGTCGAGATTGGTCTTGCCAAACAGGGGCTGAGCTGTACCTTGTAAATACACT
D N G Q H M W R P K R F R P V Y C N L C E S S I G L G K Q G L S C N L C K Y T

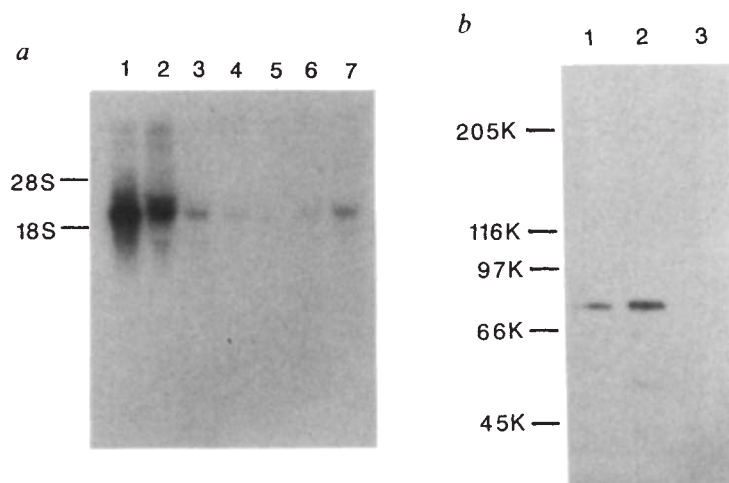
721
241

GTTCATGACCACTGTGCCATGAAGGCCCTGCCCTGTGAAGTCAGCACTATGCGAAGTCTGGGAAGGACATTGGTGTCCAAACAGCATGTGTGGTTCGAGAGGCTGTGAATCTTGGGGC
V H D Q C A M K A L P C E V S T Y A K S R K D I G V Q T H V W V R G G C E S G R

CACAGTCTGGTGGCGCTGCATTGTGTGTGGTGCACCTAGAGATCCATGATGACTGCTACCAGCCATGGGCCATGAGTGTGACTGC
 H S L V L G L H C V W C H L E I H D D C L P A M G H E C D C
 TCCATCTATCCCAAGTGTCTGGCCTCTGGACAGGAACGTAAAGTTAGCAAAAACAAGCCAGAACCCACAGATGTTTAAATTTGAGC
 S I Y P S V L A S G Q R E K V S K T S Q K T T D D L N L S
 AACACGCACCCACTTCTGGTCTTTGTCAACCCCAAGAGTGGCGGAAGCAAGGGGAGAGGGTCTTTGAAATTCAGTATCTACTA
 N T H P L L V F V N P K S G G K Q G G E R V L W K F Q Y L L
 GGTCTGAGCAGGGCTCAGATTCTTCAGAGAGTTCCTGATTACGGGATTTTGGTGTGTGGCGAGATGGCAGTAGGCTGGATC
 G P E P G L R F F R E V P D Y R I L V C G G D G T V G W I
 GTGCTCTCTGTTGCGGTGTTCGCTGGGCACTGGGCACTGAAATGCTATGCTGCTCAGATGGGAGGAGTTATGAGGGACAGAA
 V P P V A V L L P L G T G N D L R T G C L R W G G G Y E G Q N
 AAGGTGGTACACATGGATCGATGTGCTGGAGGTGATACCCCAACAACCTAGAAAGAGGACGCCAGCTCCCTTCATGATCATC
 K V V H M D R W S V E V I P Q Q T E E K S D P V P F Q I I
 TATGGCTCACGATTCACATCATCGCAGAGAAATATCTCGAGAAGTTCACAGCAGAATGAAGAAACAAGCTGTGGTACTTCGAATTT
 I A H R F H I M R E K Y P E K F N S R M K N K L W Y P F F
 AAATCTGGAGGAATTTGACATGTCAGATGTGTGGGAACCAATGGATCTGAGCAACCTGTCCCTAGAAAGTATCGCAGATGCTGAAC
 K L E E S L T V E I C G K P P L D L S N L S L E G I A V L N
 GGGTGAATACCAAGAGACCGCATGGGATATCCACGGGATACCAACGGGCTTAGCGGCATGGCCAAAGTCAACCGCACCCOGATATC
 G D T K R P H G D I H G I N Q A L G A M A K V I T D P D I
 CGGCTCGAGGAGTAGTAGGCTGGAAGTGCAATGATGAGTGGGCGAGATCTATACCAAGATCAAGATGCTGGACATGGCTAGCCAA
 R L E V V G L E G A G I E M G Q I Y T K L N A G H R L A K
 TCCCCATGCAAAATCGATGAGAACCGGTGATGCAGACGCCCTGTACGATCAAGTACACCCACAGGAACAGATGCCCATGCTGTG
 L P M Q I D G E P W M Q T P C T I K I T H R N Q M P M L V
 TTTCTTGTGCTAGGGGGACACCTCAGGCTCAGCAAGCGGCTTGAAACCCACTCATTTGTCCTGGACTCTACTCCCTAGGCTCTGTA
 F L C end
 AGTGTAAATACCTCACAGATTTTATTATCTGCACAACCTCCACTATTTCCCATGAACCCATGACACACGACACACACACACA
 CAGATGAAGTGCTCTCTGTAATAAAATACTTTTTTTCCTACGGGATAAAAAAAA

FIG. 2 Northern hybridization analysis of porcine tissues and detection of protein expressed in the transformant. *a*, Northern hybridization of RNA from porcine tissues using a cDNA probe (λ DGK1). Except for lane 1 (poly(A)⁺ RNA from thymus), total RNAs from thymus (lane 2), spleen (3), kidney (4), liver (5), heart (6) and brain (7) were analysed. Size markers are 28S and 18S porcine ribosomal RNAs. *b*, Detection of translation product of λ DGK1 clone by Western blotting. 80K DGK from porcine thymus was analysed as a standard (lane 1, 90 ng of protein). The extracts from *E. coli* transfected with λ DGK1 (lane 2, 4.5 μ g) and from cells transfected with the vector alone (lane 3, 4.5 μ g) were analysed by western blotting using anti-80K DGK antibody.

METHODS. *a*, Total RNA (20 μ g) or poly(A)⁺ RNA (10 μ g) was denatured with 1 M glyoxal and 50% dimethylsulphoxide²⁸, electrophoresed on 1% agarose, and transferred to a nylon membrane (Nytran, Schleicher and Schuell)²⁸. The hybridization probe was λ DGK1 labelled with ³²P (Amersham) by a multipriming procedure²⁹. Otherwise hybridization conditions were as in ref. 24. *b*, As the λ DGK1 clone was not properly ligated into the reading frame of λ gt11 vector, the clone was inserted into pBluescript for frame adjustment. The *E. coli* cells (XL1-Blue, Stratagene) transfected with the λ DGK1 clone or with vector alone were cultured for 12 h, and the fusion protein was then induced by further incubation for 3 h with isopropyl- β -D-thiogalactopyranoside (5 mM; Sigma). Cells were then resuspended in Laemmli sample buffer³⁰, boiled for 3 min, spun to remove insoluble material



and the extract electrophoresed in a 7.5% SDS-polyacrylamide gel before transferring to a nitrocellulose membrane (Schleicher and Schuell). Detection with the anti-80K DGK antibody is described in ref. 2.

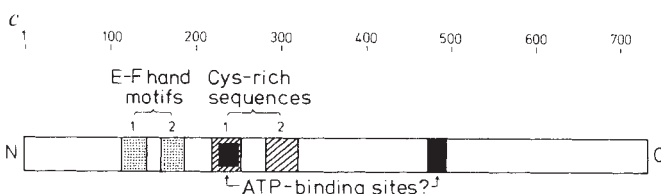
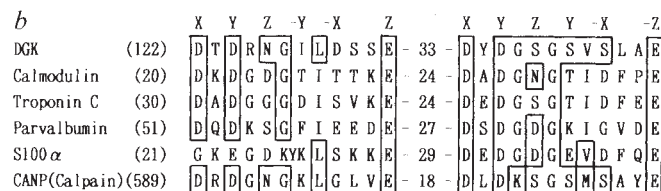
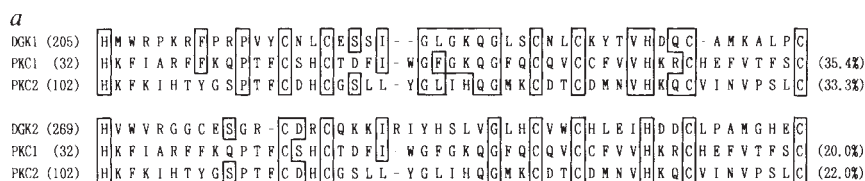
activity is thus apparently regulated by both Ca²⁺ and ionic amphiphiles.

DGK also contains two sets of the sequence Gly-X-Gly-X-X-Gly---Lys at residues 226–248 (within the first cysteine-rich sequence) and 473–492. This sequence is found in protein kinases and is thought to constitute the ATP-binding site¹⁹, but it is not clear whether this is the case in DGK, a lipid kinase. Also, the significance of the sequence location within the first cysteine-rich segment is not obvious. Apart from these putative ATP-binding sites, no significant similarities can be found between the DGK sequence and the catalytic domains of other kinases sequenced so far. The sequence of *E. coli* DGK²⁰ does

not show any significant similarity to the 80K DGK either. The 80K DGK is phosphorylated on serine by protein kinase C (ref. 21) and in fact contains several serine residues (for example, at positions 2, 23, 153, 278, 491 and 649) as potential phosphorylation sites²².

The structure of porcine 80K DGK is summarized schematically in Fig. 3c. The primary structure deduced for porcine 80K DGK therefore contains two cysteine-rich, zinc finger-like sequences, like the PKCs, and two sets of E-F hand motifs as potential Ca²⁺-binding sites. Although the function of the cysteine-rich sequences is still unknown, the finding that the enzyme activity is stimulated by Ca²⁺ in the presence of an anionic or

FIG. 3 Comparison of amino-acid sequences. *a*, Amino-acid sequences of cysteine-rich, zinc finger-like regions of DGK and PKC. Rat PKC α type was taken as a representative of the PKC family. Identical amino acids are boxed. The residue number of the first amino acid is given in parenthesis at the end of each column. DGK1 and DGK2, and PKC1 and PKC2 indicate the first and second cysteine-rich sequences of DGK and PKC, respectively. *b*, Comparison of amino-acid sequences of Ca²⁺-binding loops in E-F hand motifs between DGK and several Ca²⁺-binding proteins. Identical amino acids are boxed. Numbers in parentheses denote the residue number of the first amino acid of each protein. Numbers between amino-acid sequences represent amino-acid residues between the Ca²⁺-binding loops. X, Y, Z, -Y, -X and -Z denote the calcium-coordinating positions¹⁷. Proteins used for comparison are chicken calmodulin³¹, rabbit troponin C³², carp parvalbumin¹⁷, bovine S-100 protein (α -subunit, S100 (ref. 33)) and the large subunit of rabbit calcium protease¹⁶ (μ -type, CANP or calpain I). *c*, Schematic presentation of the structure of 80K DGK. N, Amino terminus; C, carboxyl terminus.



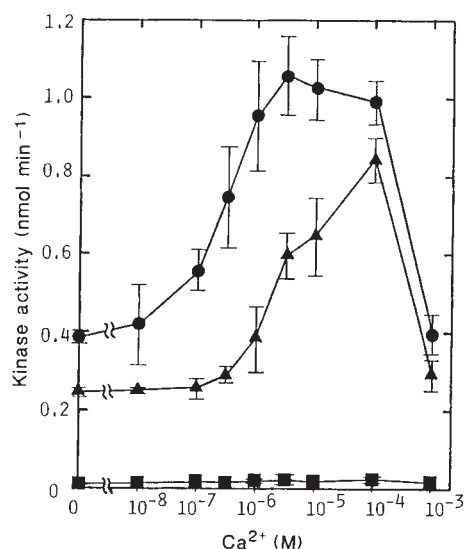


FIG. 4 Effect of calcium on the activity of 80K DGK in the presence of 1 mM deoxycholate (●) or 10 μ M sphingosine (▲) as activator; (■) control without activator.

METHODS. DGK activity was assayed using purified 80K DGK (50 ng protein) as described in ref. 3, except that MOPS (50 mM, pH 7.2) was used instead of Tris as buffer. The reaction mixture contained 1.0 mM [γ - 32 P]ATP and 0.5 mM 1,2-dioleoylglycerol. Before assay, the enzyme was treated with 10 mM EGTA at 0 °C for 30 min to chelate endogenous calcium. The free calcium concentrations indicated were achieved by varying the calcium/EGTA ratio³⁴. EGTA concentration was kept at 1.0 mM in the assay. Results are averages from three incubations.

cationic amphiphile indicates that the E-F hand motifs are important. DGK regulation by Ca^{2+} may be different from the PKCs, which have no obvious Ca^{2+} -binding sites^{8,23}; regulation could involve inositol 1,4,5-trisphosphate, another second messenger that induces Ca^{2+} mobilization in stimulated cells¹¹. Several immunologically distinct DGK isoforms have been reported¹⁻³, and we have recently purified a DGK isoenzyme which, in contrast to the 80K enzyme, does not require lipid activators for its activity³. Sequence analysis of these DGK isoenzymes should clarify the structure-function relationships in this enzyme family. □

Received 13 November 1989; accepted 16 January 1990.

- Kanoh, H., Yamada, K. & Sakane, F. *Trends Biochem. Sci.* **15**, 47-50 (1990).
- Yamada, K. & Kanoh, H. *Biochem. J.* **255**, 601-608 (1988).
- Sakane, F., Yamada, K. & Kanoh, H. *FEBS Lett.* **255**, 409-413 (1989).
- MacDonald, M. L., Mack, K. F., Williams, B. W., King, W. C. & Glomset, J. A. *J. Biol. Chem.* **263**, 1584-1592 (1988).
- Yamada, K., Sakane, F. & Kanoh, H. *FEBS Lett.* **244**, 402-406 (1989).
- Kanoh, H., Kondoh, H. & Ono, T. *J. Biol. Chem.* **258**, 1767-1774 (1983).
- Kanoh, H., Iwata, T., Ono, T. & Suzuki, T. *J. Biol. Chem.* **261**, 5597-5602 (1986).
- Nishizuka, Y. *Nature* **334**, 661-665 (1988).
- Kretsinger, R. H. *Adv. Cyclic Nucleotide Res.* **11**, 1-26 (1979).
- Nishizuka, Y. *Nature* **308**, 693-697 (1984).
- Berridge, M. J. & Irvine, R. F. *Nature* **341**, 197-205 (1989).
- Kozak, M. *Cell* **44**, 283-292 (1986).
- Berg, J. *Science* **232**, 485-487 (1986).
- Ono, Y. et al. *Proc. natn. Acad. Sci. U.S.A.* **86**, 4868-4871 (1989).
- Csermely, P., Szamel, M., Resch, K. & Somogyi, J. *J. Biol. Chem.* **263**, 6487-6490 (1988).
- Emori, Y. et al. *J. Biol. Chem.* **261**, 9465-9471 (1986).
- Tufty, R. H. & Kretsinger, R. H. *Science* **187**, 167-169 (1975).
- Chou, P. Y. & Fasman, G. D. *Rev. Biochem. Chem.* **47**, 251-276 (1978).
- Hunter, T. *Cell* **50**, 823-829 (1987).
- Loomis, C. R., Walsh, J. P. & Bell, R. M. *J. Biol. Chem.* **260**, 4091-4097 (1985).
- Kanoh, H., Yamada, K., Sakane, F. & Imaizumi, T. *Biochem. J.* **258**, 455-462 (1989).
- Kishimoto, A. et al. *J. Biol. Chem.* **260**, 12492-12499 (1985).
- Ohno, S. et al. *Nature* **325**, 161-166 (1987).
- Takai, T., Yokoyama, C., Wada, K. & Tanabe, T. *J. Biol. Chem.* **263**, 2651-2657 (1988).
- Gubler, U. & Hoffman, B. J. *Gene* **25**, 263-269 (1983).
- Yanisch-Perron, C., Vieira, J. & Messing, J. *Gene* **33**, 103-119 (1985).

- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
- Feinberg, A. P. & Vogelstein, B. *Analyt. Biochem.* **132**, 6-13 (1983).
- Laemmli, U.K. *Nature* **227**, 680-685 (1970).
- Putney, J. A. et al. *J. Biol. Chem.* **258**, 11864-11870 (1983).
- Wilkinson, J. M. *Eur. J. Biochem.* **103**, 179-188 (1980).
- Kuwano, R. et al. *FEBS Lett.* **202**, 97-101 (1986).
- Burgess, G. M., McKinney, J. S., Fabiato, A., Leslie, B. A. & Putney Jr J. W. *J. Biol. Chem.* **258**, 15336-15345 (1983).

ACKNOWLEDGEMENTS. We thank S. Imai for helping us to sequence the cDNAs; Professor R. Sato of the Institute for Protein Research, Osaka University, for encouragement and critical reading of the manuscript, and Professor K. Fujinaga of the Cancer Research Institute, Sapporo Medical College, for the use of facilities in his laboratory. Supported in part by the Naito Foundation, Japan.

Enhanced myogenesis in NCAM-transfected mouse myoblasts

George Dickson, Dave Peck, Stephen E. Moore, C. Howard Barton & Frank S. Walsh

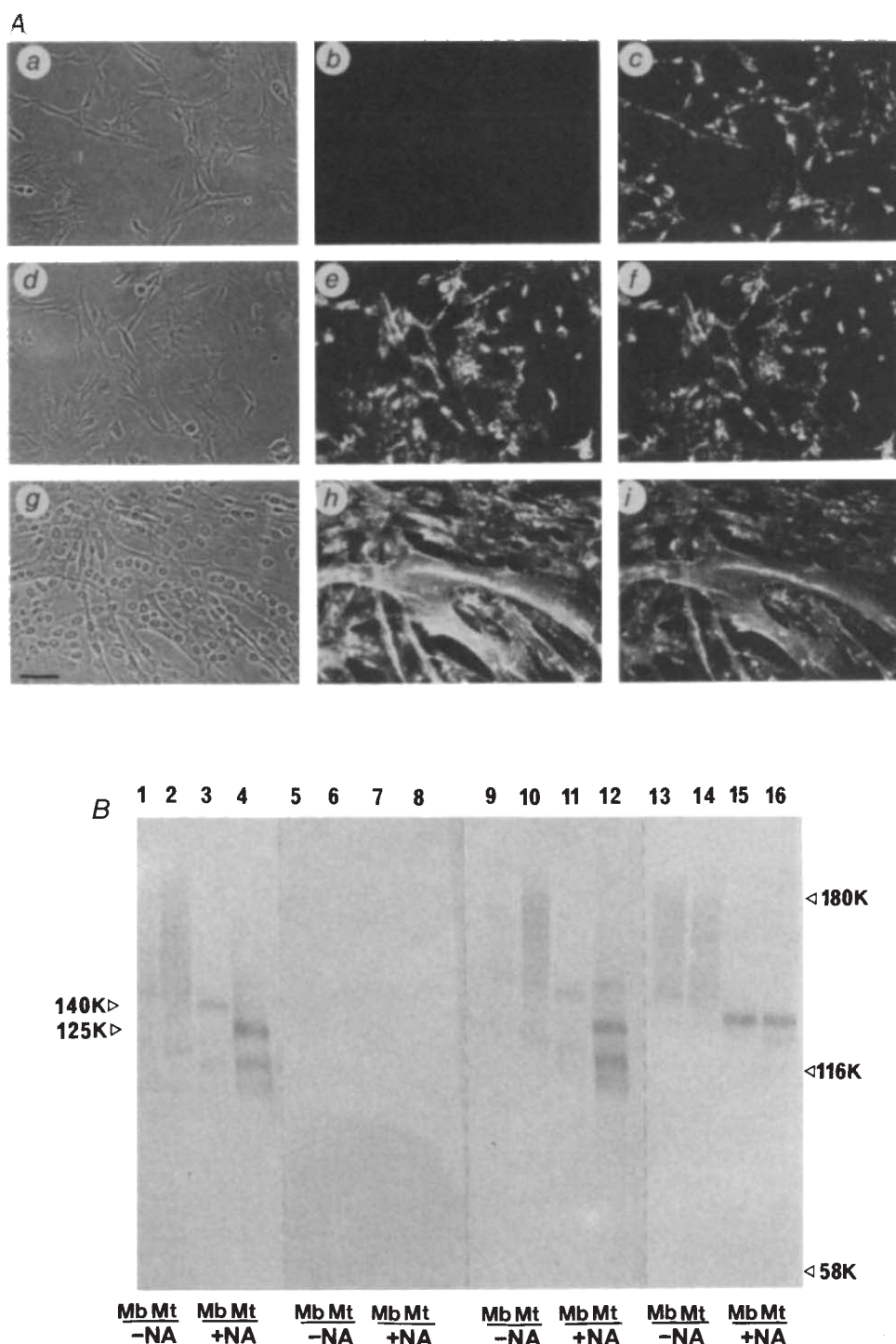
Department of Experimental Pathology,
United Medical and Dental Schools, Guy's Hospital,
London Bridge, London SE1 9RT, UK

THE fusion of mononucleate precursor myoblasts to form the multinucleated skeletal muscle fibre is preceded by a series of complex cell-cell interactions¹⁻⁵ but the cell-surface molecules involved in these events have not been characterized. During myogenesis *in vivo* and *in vitro*, expression of the neural cell adhesion molecule (NCAM) undergoes an isoform transition that precisely correlates with terminal myoblast differentiation and myotube formation^{6,7}. Altered processing of RNA results in the replacement of the transmembrane NCAM (relative molecular mass, 145,000 (145K)) in proliferating myoblasts by a predominant 125K NCAM form linked to glycosyl phosphatidylinositol in myotubes^{8,9}. We now report that mouse myoblasts transfected to constitutively express the human muscle-specific 125K glycosyl-phosphatidylinositol-linked NCAM isoform more readily fuse to form myotubes. This suggests that NCAM plays a part in myoblast fusion and that the isoform switch may promote this function.

The process of terminal commitment and differentiation that leads to myotube formation involves the withdrawal of proliferating myoblasts from the cell cycle and activation of a battery of muscle gene products, with concurrent changes in myoblast cell surface and adhesive properties^{1,2,10}. Specific inhibitors of myogenesis have been widely used to define discrete components of this multistep process and to identify essential stages in the progression to myotube formation^{2,11}. Fusion will not occur in the absence of specific gene activation nor of a perturbed cytoskeleton, and studies on Ca^{2+} withdrawal indicate that specific Ca^{2+} -dependent cell interactions are an essential component. The contribution of other cell adhesion systems in the process of myoblast-myoblast recognition, aggregation and membrane fusion cannot be excluded however. For example, immunoblockade of integrin-fibronectin interaction inhibits fusion by maintaining cells in a state of proliferation¹², but it has no direct effect on fusion *per se*⁴. The pattern of NCAM expression has been found to be tightly regulated during development and regeneration of skeletal muscle, which suggests that it has a role in myoblast fusion and nerve-muscle interaction^{6,7,13,14}, although only the latter has been demonstrated^{13,15-17}. A small (20%) but non-significant inhibition of myotube formation in chick muscle cultures by antibodies to embryonic brain NCAM^{13,17} has been described, and recently it has been suggested that the inhibitory effects of phosphatidylinositol-specific

FIG. 1 Expression of human NCAM in transfected mouse C2 myoblasts. **A**, Cell-surface immunostaining of parental C2 myoblasts (*a-c*), and of myoblast (*d-f*) and myotube (*g-i*) cultures of a C2 cell line (NCAM-B11) transfected with a gene construct encoding the 125K GptdIns-linked human muscle NCAM⁹. Cells shown in phase micrographs (*a, d, g*) were double-labelled with species-specific antibodies for human (*b, e, h*) and mouse (*c, f, i*) NCAM and visualized with fluorescein- and rhodamine-conjugated second antibodies respectively. Scale bar, 100 μ m. **B**, Western immunoblots of parental C2 (tracks 1-8) and human NCAM-transfected cells (tracks 9-16) using species-specific antibodies for mouse (tracks 1-4 and 9-12) and human NCAM (tracks 5-8 and 13-16). Extracts of myoblast (mb) and myotube (Mt) cultures were analysed with (+NA) or without (-NA) prior treatment with neuraminidase to remove polysialic acid side chains and to resolve core desialo polypeptides. Migration of M_r standards is indicated.

METHODS. C2 cells were transfected with the human NCAM gene construct, and G418-resistant clones screened by immunostaining with anti-human NCAM monoclonal antibody as previously described²⁸. Dual-label immunostaining was performed in conjunction with a rabbit antiserum to mouse NCAM⁷. To induce fusion, proliferating myoblasts were collected and reseeded at 3×10^5 cells per cm^2 in DME medium containing 2% horse and 0.5% fetal calf serum. Myotubes were analysed 4 days later. Western immunoblotting after SDS-PAGE has been described using rabbit anti-mouse NCAM^{7,19} or mouse anti-human NCAM antibodies²⁰.



phospholipase C on myotube formation result from cleavage of cell-surface NCAM¹⁸. As an alternative investigation of the possible function of NCAM in the events leading to myoblast fusion, we have examined myogenesis in conditional NCAM mutants produced by transfection-based overexpression of human NCAM in the mouse myoblast cell line C2 (ref. 19).

C2 myoblasts were stably transfected with a complementary DNA gene construct encoding the 125K glycosylphosphatidylinositol (GPI)-tailed human NCAM^{9,20,21} and positive clones were identified by immunostaining with a species-specific anti-human NCAM monoclonal antibody^{20,22,23}. As seen from Fig. 1A, introduction of the human NCAM causes it to be

coexpressed on the surface of C2 cells with the endogenous mouse gene product; it is present in both myoblasts and myotubes. Also, western blot analysis of transfected cells (Fig. 1B) shows that a single human NCAM species (125K) (is present as well as the normal endogenous mouse forms⁷.

A clear enhancement of fusion is evident in the light microscope in several transfected lines by comparison with the C2 control cells. The effect is striking at early times (2-3 days) after the start of myogenic differentiation but cultures are undistinguishable at later stages (Fig. 2A). We quantitated the extent of myogenesis by determining the per cent of nuclei in myotubes and the creatine phosphokinase activity: 2 days after initiation

FIG. 2 Time-course of myogenesis in parental and human NCAM-transfected C2 myoblasts. **A**, Cultures of C2 myoblasts (*a-c*) and transfected NCAM-B11 cells (*d-f*) were induced to form myotubes, and Giemsa-stained cultures were examined under bright-field illumination after 1 day (*a, d*), 2 days (*b, e*) and 4 days (*c, f*) under fusion-promoting conditions. Scale bar, 100 μm . **B**, Qualitative morphological assessment of myoblast fusion (% of nuclei in myotubes) in parental C2 ($\circ$) and transfectant NCAM-B11 ($\bullet$) cultures. Total nuclear counts for C2 ($\square$) and NCAM-B11 ($\blacksquare$) are given as a measure of cell number. **C**, Assessment of CPK-specific activity as a parameter of biochemical differentiation in fusing C2 ($\circ$) and NCAM-B11 transfectant cells ($\bullet$). Values in **B** and **C** are means $\pm$ s.e.m. ($n=3$) of independent cultures.

METHODS. Myoblast fusion was induced as described in the legend to Fig. 1. For morphological analysis, cultures were fixed with 95% ethanol/5% acetic acid (5 min; -20°C) and methanol (15 min; 20°C), stained with 20% Giemsa (15 min, 20°C) and the total number of nuclei and myotube-associated nuclei counted¹ in four fixed-position fields of 1 mm^2 using an eyepiece graticule and $\times 200$ magnification. For CPK determinations, cultures were collected in 15 mM NaCl, 1 mM EDTA, pH 7.5, sonicated and assayed for CPK activity (BCL; UV-NACS Kit) and protein (Bio-Rad Bradford reagent).

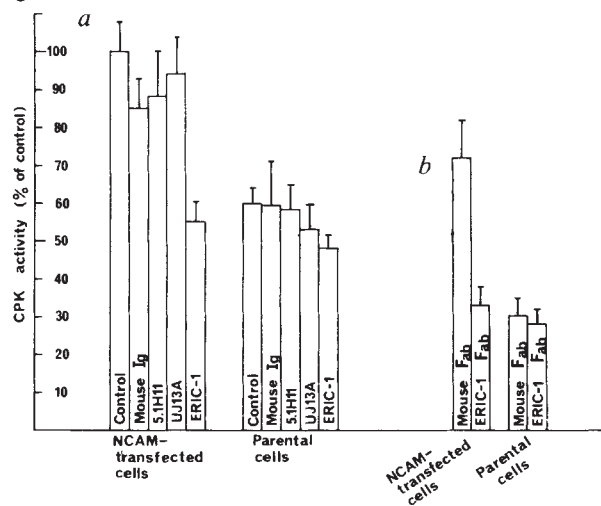
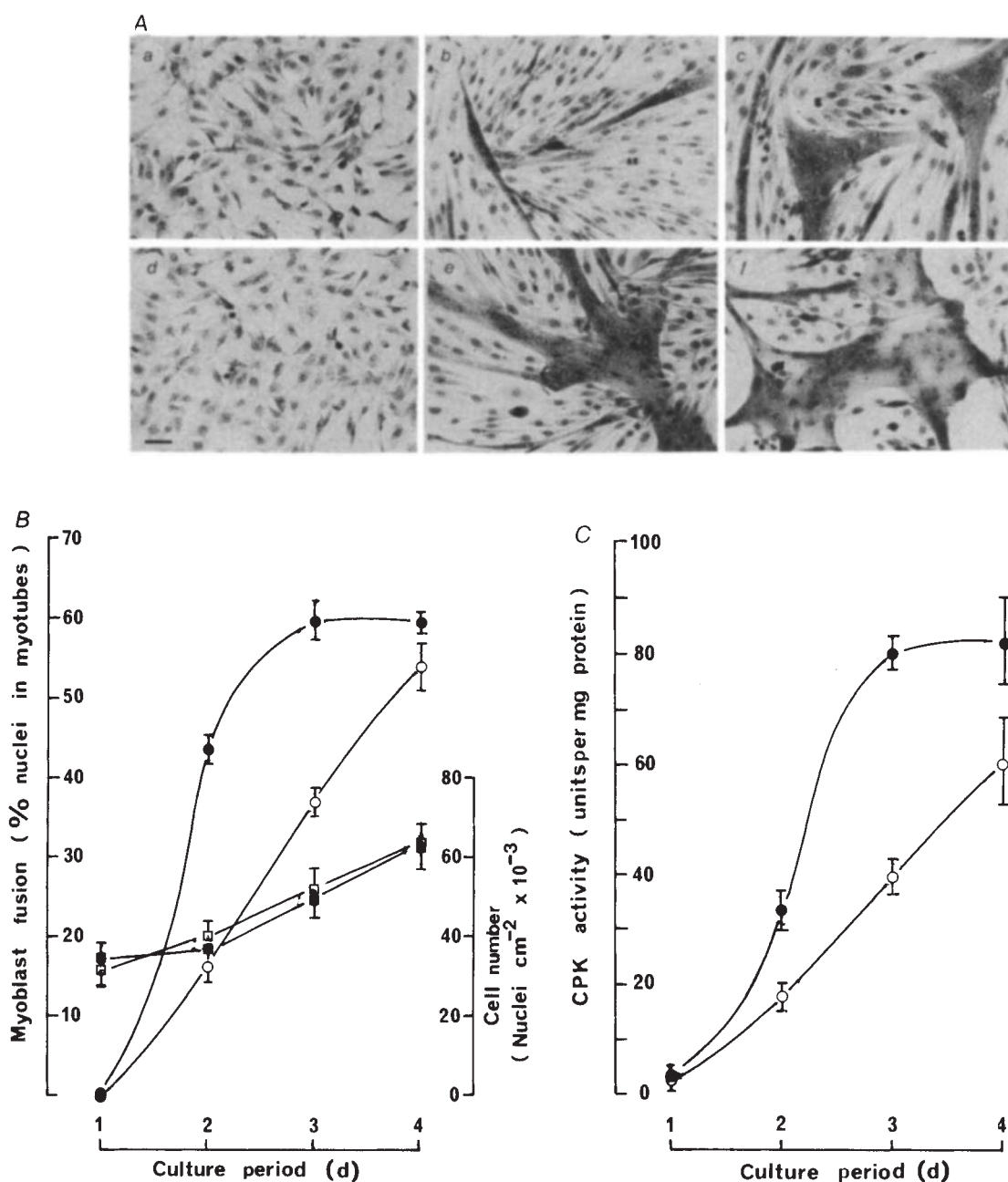


FIG. 3 Antibodies to human NCAM specifically inhibit myogenesis in transfected C2 cells. Cultures of parental and NCAM-transfected (NCAM-B11) C2 cells were induced to fuse in the presence of *a*, 300 $\mu\text{g ml}^{-1}$ immunoglobulin of non-immune mouse serum, or of anti-human NCAM monoclonal antibodies 5.1H11, UJ13A and ERIC-1, and *b*, 200 $\mu\text{g ml}^{-1}$ Fab' fragments from mouse control or ERIC-1 antibody. After 2 days, CPK was assayed as described in the legend to Fig. 2: results are shown as $\% \pm$ s.e.m. ($n=3$) of the value obtained in control untreated NCAM-B11 cells.

METHODS. All three antibodies are IgG subtypes, and IgG and derived monovalent Fab fragments were prepared by standard procedures using agarose-conjugated protein A or protein G, and immobilized papain (Pierce)²⁹. Immunoglobulin and Fab preparations were dialysed into DME medium before their addition to cultures for the culture period.

TABLE 1 Comparison of cell-surface expression and enhancement of myogenesis in NCAM-transfected C2 myoblasts

	Human NCAM cell-surface expression (Units* per 10 ⁶ -cells)	Myoblast fusion (% nuclei in myotubes)	CPK activity (units mg ⁻¹)	Cell number (nuclei per cm ² × 10 ⁻³)
NCAM transfectants				
NCAM-B6	47 ± 1.8	24 ± 0.6	ND	34 ± 1.3
NCAM-B9	67 ± 1.9	34 ± 2.9†	22 ± 2.4‡	35 ± 3.7
NCAM-B4	74 ± 2.4	36 ± 5.8‡	23 ± 1.1†	37 ± 7.0
NCAM-B10	159 ± 2.3	48 ± 2.0§	28 ± 2.7†	40 ± 2.9
NCAM-B11	187 ± 3.5	56 ± 1.8§	30 ± 1.1§	39 ± 2.0
Parental cells				
C2	0.58 ± 0.1	22 ± 0.9	14 ± 0.9	42 ± 5.6
Vector-alone transfectants				
Neo-W11	ND	19 ± 0.2	18 ± 1.2	38 ± 7.0
Neo-W12	ND	22 ± 3.1	14 ± 1.8	37 ± 1.8
Neo-W13	ND	15 ± 1.4	15 ± 0.3	35 ± 1.2

A number of independent C2 clones were isolated after transfection with the human NCAM gene construct (NCAM-B4, B6, B9, B10 and B11) or with the wild-type expression vector (Neo-W11, W12 and W13). Using an enzyme-linked immunosorbent assay (ELISA)²⁰, the relative cell-surface expression of human NCAM was determined; cultures were analysed for myoblast fusion, creatine phosphokinase (CPK) activity and total cell number after 2 days under fusion-promoting conditions (for methods see legend to Fig. 2). Values are means ± s.e.m. ($n=3$) for independent cultures and statistical significance values (†, $P<0.01$; ‡, $P<0.05$; §, $P<0.001$) were calculated using the Student's t -test to compare NCAM transfected lines with parental C2 cells. ND, Not determined.

* Cells (1,000–5,000) in microtitre plate wells were assayed by ELISA within the linear response range of the assay²⁰. Units represent absolute optical density values adjusted for cell number.

|| The human NCAM value for C2 cells represents the background reading in the ELISA system.

of fusion, myotube formation and differentiation increases significantly in transfected cells compared with C2 controls (Fig. 2B). Furthermore, as shown in Table 1, increased cell-surface expression of human NCAM in different clonal lines of transfected C2 cells correlates with an enhancement of myogenesis, as assayed by myotube formation and creatine phosphokinase activity.

Finally, to assess the contribution of the transfected human NCAM in this enhancement of fusion, we tested species-specific anti-human NCAM monoclonal antibodies (5.1H11, UJ13A and ERIC-1; refs 22, 23) for immunoblocking activity. As shown in Fig. 3, addition of 5.1H11, UJ13A or mouse control immunoglobulins causes small, non-significant fluctuations in myogenesis of both parental C2 and transfected cell lines. By contrast, ERIC-1 antibody and its derived monovalent Fab fragments completely reverse enhanced myogenesis in the transfected cells, but have no effect on C2 parental cells; this indicates that the neutralization of the capacity of the human gene product to enhance the rate of myoblast fusion is species-specific.

Immunoneutralization^{13–16,24} and transfection methods^{20,21,25} have been used to define the role of NCAM in cellular events: the function of the different isoforms in cell-cell aggregation and in neurite outgrowth has been investigated using NCAM-

negative host cells^{20,21,25}. Our strategy here has been to establish whether overproduction of exogenous NCAM alongside the endogenous forms that are normally regulated perturbs myotube formation. We find that constitutive high-level expression of the human muscle-specific 125K glycosylphosphatidylinositol-linked NCAM isoform in mouse C2 cells enhances the rate of myoblast fusion and differentiation. As this isoform is normally activated during commitment of myoblasts to terminal differentiation^{6,7}, the effect of NCAM transfection may be to prematurely mimic one of the naturally occurring regulatory events that accompany myogenesis. Factors other than NCAM that are involved in the complex events leading to myoblast fusion include Ca²⁺-dependent cell-cell recognition events^{3,4}, interaction with the extracellular matrix¹² and availability of certain polypeptide growth factors^{26,27}. In particular, NCAM-NCAM interaction could be the Ca²⁺-independent component that complements and/or triggers the essential Ca²⁺-dependent cell-cell interaction steps of myoblast fusion¹⁸. As both fusion and creatine phosphokinase activation are enhanced in NCAM-transfected myoblasts, NCAM may mediate not only cell-surface recognition and adhesion, but could also affect the programme of gene activation that accompanies terminal myoblast differentiation. □

Received 25 September 1989; accepted 24 January 1990.

1. Bischoff, R. in *Cell Surface Reviews, Membrane Fusion* Vol. 5 (eds Poste, G. & Nicholson, G. L.) 127–175 (North-Holland, Oxford, 1978).
2. Wakelam, M. J. O. *Biochem. J.* **228**, 1–12 (1985).
3. Gibraltier, D. & Turner, D. C. *Dev. Biol.* **112**, 292–307 (1985).
4. Knudsen, K. A. *J. Cell Biol.* **101**, 891–897 (1985).
5. Pizze, J. A., Jones, G. E. & Walsh, F. S. *J. Cell Biol.* **107**, 2307–2317 (1988).
6. Covault, J., Merlie, J. P., Goridis, C. & Sanes, J. R. *J. Cell Biol.* **102**, 731–739 (1986).
7. Moore, S. E., Moore, S. E., Thomson, J., Kirkness, V., Dickson, G. & Walsh, F. S. *J. Cell Biol.* **105**, 1377–1386 (1987).
8. Dickson, G. *et al. Cell* **50**, 1119–1130 (1987).
9. Thomson, J. *et al. Genes Dev.* **3**, 348–357 (1989).
10. Knudsen, K. A. & Horwitz, A. F. *Dev. Biol.* **58**, 329–338 (1977).
11. Endo, T. & Nadal-Ginard, B. *Cell* **49**, 515–526 (1987).
12. Merko, S. A. & Boettiger, D. *Cell* **51**, 51–57 (1987).
13. Rutishauser, U., Grumet, M. & Edelman, G. M. *J. Cell Biol.* **97**, 145–152 (1983).
14. Bixby, J. L., Pratt, R. S., Lilien, J. & Reichardt, L. F. *Proc. natn. Acad. Sci. U.S.A.* **84**, 2555–2559 (1987).
15. Rieger, F. *et al. J. Cell Biol.* **107**, 707–719 (1988).
16. Landmesser, L., Dahm, L., Schultz, K. & Rutishauser, U. *Dev. Biol.* **130**, 645–670 (1988).
17. Grumet, G. M., Rutishauser, U. & Edelman, G. M. *Nature* **295**, 693–695 (1982).

18. Knudsen, K. A., Smith, L. & McElwee, S. *J. Cell Biol.* **109**, 1779–1786 (1989).
19. Yaffe, D. & Saxel, O. *Nature* **270**, 725–727 (1977).
20. Doherty, P. *et al. J. Cell Biol.* **109**, 789–798 (1989).
21. Pizze, J., Rowett, L. H., Barton, C. H., Dickson, G. & Walsh, F. S. *J. Cell Biol.* **109**, 789–798 (1989).
22. Walsh, F. S., Dickson, G., Moore, S. E. & Barton, C. H. *Nature* **339**, 516 (1989).
23. Patel, K., Rossell, R. J., Moore, S. E., Walsh, F. S. & Kemshead, J. T. *Int. J. Cancer* **44**, 1062–1068 (1989).
24. Fraser, S. E., Murray, B. A., Chuong, C.-M. & Edelman, G. M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4222–4226 (1984).
25. Edelman, G. M., Murray, B. A., Mege, R.-M., Cunningham, B. A. & Gallin, W. J. *Proc. natn. Acad. Sci. U.S.A.* **84**, 8502–8506 (1987).
26. Massague, J., Cheifetz, S., Endo, T. & Nadal-Ginard, B. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8206–8210 (1986).
27. Lathrop, B. K., Thomas, K. & Fraser, L. *J. Cell Biol.* **101**, 2194–2198 (1985).
28. Gower, H. J. *et al. Cell* **55**, 955–964 (1988).
29. Sudher, P. *et al. Science* **244**, 1158–1162 (1989).

ACKNOWLEDGEMENTS. This work was supported by the Muscular Dystrophy Group of Great Britain, the Wellcome Trust and the Brain Research Trust.

Disruption of the actin cytoskeleton in yeast capping protein mutants

James F. Amatruda*, John F. Cannon†, Kelly Tatchell‡, Christopher Hug* & John A. Cooper*

* Department of Cell Biology and Physiology, Washington University School of Medicine, St Louis, Missouri 63110, USA

† Department of Microbiology, University of Missouri, Columbia, Missouri 65211, USA

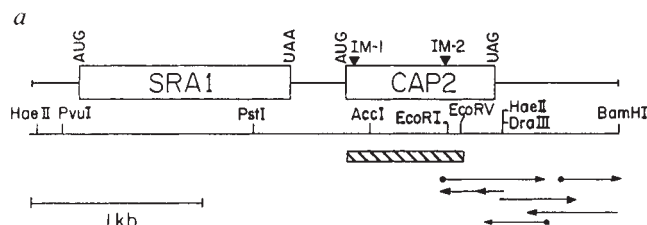
‡ Department of Microbiology, North Carolina State University, Raleigh, North Carolina 27695, USA

CAPPING protein controls the addition of actin subunits to the barbed end of actin filaments and nucleates actin polymerization *in vitro*. Capping protein has been identified in all eukaryotic cells examined so far; it is a heterodimer with subunits of relative molecular masses 32,000–36,000 (α -subunit) and 28,000–32,000 (β -subunit)^{1,2}. In skeletal muscle, capping protein (CapZ) probably binds the barbed ends of actin filaments at the Z line³. The *in vivo* role of this protein in non-muscle cells is not known. We report here the characterization of CAP2, the single gene encoding the β -subunit of capping protein in *Saccharomyces cerevisiae*. Yeast cells in which the CAP2 gene was disrupted by an insertion or a deletion had an abnormal actin distribution, including the loss of actin cables. The mutant cells were round and large, with a heterogeneous size distribution, and, although viable, grew more slowly than congenic wild-type cells. Chitin, a cell wall component

restricted to the mother–bud junction in wild-type budding yeast, was found on the entire mother cell surface in the mutants. The phenotype of CAP2 disruption resembled that of temperature-sensitive mutations in the yeast actin gene ACT1 (ref. 4), indicating that capping protein regulates actin-filament distribution *in vivo*.

Figure 1 shows the physical map, complete nucleotide sequence and predicted corresponding polypeptide sequence of the CAP2 gene, which had been previously cloned and partially sequenced⁵. CAP2 was identified as the gene encoding the β -subunit of capping protein by its similarity to the β -subunit of chicken skeletal muscle capping protein (CapZ β)⁶ shown in Fig. 2; also shown is the alignment of CAP2 with the β -subunit of capping protein from *Dictyostelium discoideum* (Cap32)⁷. The amino-acid sequences of the yeast and chicken proteins are 49% identical with eight gaps, and the amino-acid sequences of the yeast and *Dictyostelium* proteins are 43% identical with 10 gaps. These values are higher than those reported for yeast and equine tropomyosins (~20%)⁸, and yeast and bovine profilins (29%)⁹.

Capping protein blocks the barbed end of actin filaments, and nucleates actin polymerization, but does not sever actin filaments^{1,2,10}. When microinjected into tissue culture cells, capping protein disrupts the actin stress fibres¹¹. We hypothesized



b

1	GTAAAGGG	GAAGCGATCT	TAAATCCCAT	GAATTTCCTT	TCTCTTTTTC	TTCTTCTTCT	60
61	GTTCCCTTTT	TCTTTCTTCT	TCCTTCAACG	TCTACGTAAA	TATAATGTAT	AACATCTCCA	120
121	CTTTCCTTTC	CCTTATAGCG	TTATGTTACC	TGTTTAAACT	CTAAATTTCT	CATCTACTAC	180
181	TGCTGTTACT	ACATGTATAT	ATGTTTCATT	TTTTTTTTTC	ACCACGGTAA	TCGTTAAAAA	240
241	TGAGAGCATT	TGTAACAAAT	AAAGAGGAAC	TGTTCAATCA	AGTGACATAC	AAAAACTCTC	300
301	TCCAACTCTA	CAATCACCAC	CGTAATGTCT	GATGCTCAAT	TCGATGCTGC	TTTAGATCTT	360
			M S D A Q F D A A L D L				
361	CTTAGGAGGC	TAAATCCTAC	CACGTTACAG	GAGAATCTAA	ACAATCTGAT	CGAATTACAA	420
	L R R L	N P T	T L Q	E N L N	N L I	E L Q	
421	CCAAATTTGG	CACAAGATT	ACTATCTTCA	GTAGACGTTT	CCCTATCCAC	ACAGAAAGAT	480
	P N L A	Q D L	L S S	V D V P	L S T	Q K D	
481	TCCGCCGATT	CAAAACGGGA	GTACTTATGC	TGCGACTATA	ATCGTGATAT	TGATTCGTTT	540
	S A D S	N R E	Y L C	C D Y N	R D I	D S F	
541	AGATCGCCTT	GGTCGAACAC	TTATTACCCA	GAACATCTCC	CAAGAGTATC	ACAAGACAGC	600
	R S P W	S N T	Y Y P	E L S P	K D L	Q D S	
601	CCCTTTCCCT	CAGCCCCCTT	AAGAAATTTG	GAGATCTTAG	CCAATGACTC	TTTCGACGTT	660
	P F P S	A P L	R K L	E I L A	N D S	F D V	
661	TACAGAGATC	TCTATTATGA	AGGCGGTATC	TCCAGCGTGT	ACCTCTGGGA	CCTCAATGAA	720
	Y R D L	Y Y E	G G I	S S V Y	L W D	L N E	
721	GAAGATTTC	ATGGGCGACG	TTATGTCAGG	GTAGTGCTTA	TCAAAACCCA	ATCAGATCAC	780
	E D F N	G H D	Y A G	V V L I	K N Q	S D H	
781	AGTAATTGGG	ACAGTATCCA	TGTTTGTGAA	GTTACAACAT	CTCCTCTCTC	TCCGACAGT	840
	S N W D	S I H	V F E	V T T S	P S S	P D S	
841	TTCAACTATA	GAGTCACCAT	TACAATCATT	CTGCATCTGG	ACAAAACAAA	AACGTACCCG	900
	F N Y R	V T T	T I I	L H L D	K T K	T D Q	
901	AATTCTCATA	TGATGTTATC	CGGGAACCTG	ACAAGACAGA	CGGAAAGGA	CATCGCTATT	960
	N S H M	M L S	G N L	T R Q T	E K D	I A I	
961	GATATGTCCC	GTCCATTAGA	TGTTATTTTC	ACATCACATG	TGCCTAAATT	GGGTTCCTTA	1020
	D M S R	P L D	V I F	T S H V	A N L	G S L	
1021	ATTGAGGACA	TTGAGTCTCA	GATGAGAAAT	CTATTGGAAA	CTGTTTATT	CGAAAGACA	1080
	I E D I	E S Q	M R N	L L E T	V Y F	E K T	
1081	AGAGATATCT	TCCACCAAC	AAAGAACGCA	GCCATTGCTT	CCTCTGCTGA	GGAGGCGAAT	1140
	R D I F	H Q T	K N A	A I A S	S A E	A N	
1141	AAAGACGCAC	AGGCAGAGGT	AATCAGAGGT	TTACAGTCTT	TATAGAAATA	TTTACATGTA	1200
	K D A Q	A E V	I R G	L Q S L	*		
1201	TCGCTGGTAT	ATCATATAAA	TACATACAGC	AGCACTAAGT	GGTAGAAAAG	CGTATCAAAA	1260
1261	AATCACTCGG	AACGAACACT	GAAGAAAAG	CGAAAATCA	AAAATATGAC	CAACGCTTGG	1320
1321	AAAATTTCAG	TAGCAATTAG	ACGAACCTCG	CAGTCTGTGA	GTAACACGGA	TTAAGGAGCT	1380
1381	TCATATAGCT	TTGATATCTG	GTGTATTGCA	ATTACCTACT	AGGAGTTTTC	TGAATAGGGA	1440
1441	CTTAAGAAGT	CAAATATACA	GACAAAATA	GGGGGTTGTA	GAAGGAATAT	TTGATTCGAA	1500
1501	CTATGAAATG	CAGGGTATGG	TCAGAGGCTC	GTGTTTATAC	GAATATCAAT	AAACAAGAA	1560
1561	CCGAGGAATA	TGGGATTAT	GAAATACTGT	AATTGATTGG	TCCACAAATA	CAAGGACTA	1620
1621	TGAAATTGAA	AATAAAGTTG	GGCGAGGAAA	ATACTCCGAG	GTTTCCCAAG	GTGTCAAAAT	1680
1681	AGACTCTAAA	GTTAAATTTG	TTATTAAGAT	GTTGAAACCA	GTTTAAAGAA	AGAAGATCAA	1740
1741	GAGAGAAATC	AAAATCTTAA	CGGATTGTGC	TAACGAAAAA	GTGCTCCCAA	CGACTTTGCC	1800
1801	ATTTCAAAAA	GATCAATATT	ACACAAATCA	AAAGGAGGAC	GTTTTAAAT	TCATCAGGCC	1860
1861	CTATATCTTT	GATCAGCCAC	ACAATGGTCA	TGCAACATA	ATTCATCTAT	TTGATATAAT	1920
1921	AAAGGATCC	1929					

FIG. 1 The CAP2 gene and its associated polypeptide. **a**, Physical map of the CAP2 gene. The cloning, partial sequencing and insertional mutagenesis of CAP2 have been described previously⁵. IM-1 and IM-2, sites where CAP2 was disrupted by insertion of the URA3 gene⁵. Horizontal arrows show the strategy used to obtain the complete sequence. **b**, Sequencing primed with a synthetic oligonucleotide. Hatched bar, extent of the CAP2 coding region deleted in the cap2-Δ1::HIS3 allele. The chromosomal structure of insertion and deletion mutants was verified by Southern-blot analysis (data not shown). Diploid strains homozygous for cap2-1::URA3 sporulated normally and gave rise to four viable spores. **b**, Nucleotide and predicted corresponding amino-acid sequences of the CAP2 gene. The sequence shown extends from the 3' end of SRA1 to the 3' BamHI site. Consensus TATA boxes (bases 101–105 and 108–112) and transcription termination²³ signal (bases 1,338–1,340, 1,359–1,362 and 1,426–1,429) are underlined. The nucleotide and protein sequences have been submitted to GenBank under the accession number M31720. kb, Kilobase.

METHODS. To obtain a mutant strain in which 90% of the CAP2 coding region is deleted, sequences flanking the CAP2 gene were inserted into the plasmid pRS303 (ref. 24). To facilitate this construction, a BamHI restriction site was created 10 base pairs (bp) 5' of the initiating ATG of the CAP2 gene using oligonucleotide-directed mutagenesis²⁵. This plasmid was used to create a cap2 deletion strain by homologous recombination.

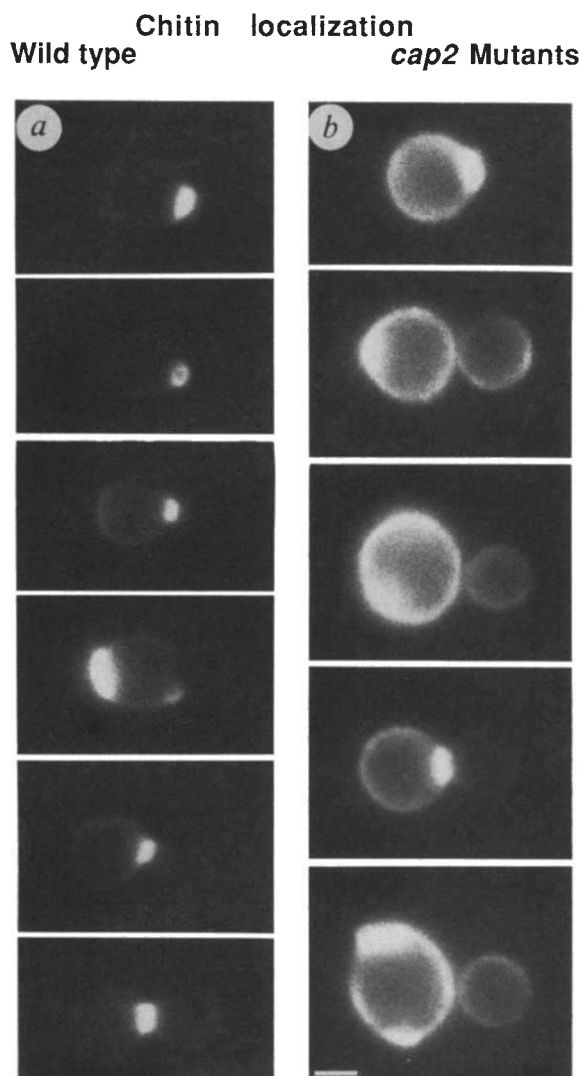


FIG. 4 Chitin distribution in wild-type and *cap2* mutant cells. Fixed cells were stained with the fluorescent dye Calcafluor and visualized with epifluorescence microscopy. *a*, Wild-type strains JC482 and YM702. *b*, *cap2* mutant strains KT607, KT610 and CH100. Scale bars, 2 μ m.

METHODS. Yeast were grown and fixed as described in the legend to Fig. 3. Cells were incubated 5 min at 23 °C in a 50 μ g ml⁻¹ solution of Calcafluor²⁹ (Fluorescent Brightener 28; Sigma), washed four times in H₂O, and processed for microscopy as described in the legend to Fig. 3.

actin cables parallel to the mother-bud axis. The bud contains spots or patches of actin predominantly at the growing tip. Shortly before cytokinesis, a few spots were seen in the mother, and a ring of actin appeared on each side of the neck. In the *cap2* mutant mother cells (Fig. 3*b*) actin cables were not present, and many actin spots were seen.

The mutant cell population had a broad distribution of sizes ranging from that of wild type (~3 μ m) to three to four times larger (~11 μ m) (Fig. 3*b*). Many of the largest cells were lysed (data not shown). The doubling time of the mutant on rich media at 30 °C was increased by 25% relative to that of the wild-type; at 39 °C the doubling time of the mutant is increased by 100% relative to the wild-type. Because deletion of the single actin gene (*ACT1*) is lethal¹⁸ the nonlethal phenotype of *cap2* mutants shows that actin has essential roles in which capping protein is not necessary. Perhaps capping protein controls the location of actin filaments, and abnormally located actin filaments are sufficient for survival.

Another abnormality in the *cap2* mutants was the delocalization of chitin deposition (Fig. 4). In wild-type cells (Fig. 4*a*),

chitin was observed in a ring at the neck of the emerging bud^{14,15}. In mutant cells (Fig. 4*b*), chitin was distributed diffusely over the entire mother cell surface and in a ring at the neck. Some mutant cells had abnormally located patches of chitin. Also, chitin rings on mutant cells were larger than those on wild-type cells.

Budding yeast grow by enlargement of the bud. Newly synthesized cell wall components are found only in the growing bud during most of the cell cycle¹⁶, indicating that this secretion is directed. Because the actin distribution in the mother and bud correlates with the site of bud emergence and growth, it has been suggested that actin contributes to the polarity of the cell and directs the location of secretion^{12,13}. The observation that *cap2* mutants have a delocalized chitin distribution as well as an abnormal actin distribution is consistent with this hypothesis. The *cap2* mutants also grew slowly, and slow growth is associated with delocalized chitin distribution¹⁷.

The phenotype of *cap2* mutants is similar to the phenotypes of mutants of actin and of certain actin-binding proteins, indicating that capping protein regulates actin *in vivo*. Strains with temperature-sensitive alleles of *ACT1* show loss of the normal asymmetric distribution of actin filaments, increased cell size and roundness, and delocalization of chitin deposition on the cell surface at the restrictive temperature⁴.

Yeast lacking profilin have a phenotype similar to that of *cap2* mutants, with the exception that, when stained with an anti-actin antibody, profilin mutants show thick bars of actin similar to those seen in strains bearing the temperature-sensitive *act1-2* allele¹⁹. Disruption of the gene encoding tropomyosin, which is localized to actin cables in wild-type cells, leads to loss of cables, changes in cell size, and to the appearance of actin bars in some cells⁸. Overexpression of an actin-binding protein of relative molecular mass 85,000 (localized to actin spots in wild type) at 37 °C has effects similar to those of *CAP2* disruption²⁰. Finally, several mutations identified as suppressors of *act1* mutants lead to disruption in cell morphology and actin distribution when present in *ACT1*⁺ cells^{21,22}. Together, these results indicate that the asymmetric distribution of actin in budding yeast depends on the proper functioning of several actin-binding proteins with apparently different functions. □

Received 1 December 1989; accepted 21 January 1990.

- Pollard, T. D. & Cooper, J. A. *Rev. Biochem.* **55**, 987–1035 (1986).
- Caldwell, J. E., Heiss, S. G., Mermall, V. & Cooper, J. A. *Biochemistry* **28**, 8506–8514 (1989).
- Casella, J. F., Maack, D. J. & Lin, S. *J. Biol. Chem.* **261**, 10915–10921 (1986).
- Novick, P. & Botstein, D. *Cell* **40**, 405–416 (1985).
- Cannon, J. F. & Tatchell, K. *Molec. cell. Biol.* **7**, 2653–2663 (1987).
- Caldwell, J. E. *et al.* *J. Biol. Chem.* **264**, 12648–12652 (1989).
- Hartmann, H. *et al.* *J. Biol. Chem.* **264**, 12639–12647 (1989).
- Liu, H. & Bretscher, A. *Cell* **57**, 233–242 (1989).
- Magdolen, V., Oechsner, U., Müller, G. & Bandlow, W. *Molec. cell. Biol.* **8**, 5108–5115 (1988).
- Cooper, J. A. & Pollard, T. D. *Biochemistry* **24**, 793–799 (1985).
- Füchtbauer, A. *et al.* *Nature* **304**, 361–364 (1983).
- Kilmartin, J. V. & Adams, A. E. M. *J. Cell Biol.* **98**, 922–933 (1984).
- Adams, A. E. M. & Pringle, J. R. *J. Cell Biol.* **98**, 934–945 (1984).
- Hayashibe, M. & Katohda, S. *J. gen. Appl. Microbiol.* **19**, 23–29 (1973).
- Cabib, E. & Bowers, B. *J. Bact.* **124**, 1586–1593 (1975).
- Novick, P. & Scheckman, R. in *Molecular Biology of the Yeast Saccharomyces: Metabolism and Gene Expression* (eds Strathern, J. N., Jones, E. W. & Broach J. R.) 361–398 (Cold Spring Harbor Laboratory, New York, 1983).
- Roberts, R., Bowers, B., Slater, M. & Cabib, E. *Molec. Cell. Biol.* **3**, 922–930 (1983).
- Shortle, D., Haber, J. E. & Botstein, D. *Science* **217**, 371–373 (1982).
- Haarer, B. K. *et al.* *J. Cell Biol.* **110**, 105–114 (1990).
- Drubin, D. G., Miller, K. G. & Botstein, D. *J. Cell Biol.* **107**, 2551–2561 (1988).
- Adams, A. E. M., Botstein, D. & Drubin, D. G. *Science* **243**, 231–233 (1989).
- Novick, P., Osmond, B. C. & Botstein, D. *Genetics* **121**, 659–674 (1989).
- Zaret, K. S. & Sherman, F. *Cell* **28**, 563–573 (1982).
- Kunkel, T. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 488–492 (1985).
- Needleman, S. B. & Wunsch, C. D. *J. molec. Biol.* **48**, 443–453 (1970).
- Devereux, J., Haeberli, P. & Smithies, O. *Nucleic Acids Res.* **12**, 387–395 (1984).
- Sherman, F., Fink, G. R. & Hicks, J. B. *Methods in Yeast Genetics* (Cold Spring Harbor Laboratory, New York, 1986).
- Sloat, B. F. & Pringle, J. R. *Science* **200**, 1171–1173 (1978).

ACKNOWLEDGEMENTS. We thank Dave Drubin for the anti-actin antibody, Mark Johnston and Liz Howell for advice and encouragement, and James Waddle and Trina Schroer for comments on the manuscript. This work was supported by the NCI (K.T.), the NIH and the Lucille P. Markey Charitable Trust. J.A.C. is a Lucille P. Markey Scholar.

Homologous activators of *ras* in fission and budding yeast

David A. Hughes*, Yasuhisa Fukui*†
& Masayuki Yamamoto*‡

* Institute of Medical Science, University of Tokyo, PO Takanawa, Tokyo 108, Japan

‡ Department of Biophysics and Biochemistry, Faculty of Science, University of Tokyo, PO Hongo, Tokyo 113, Japan

† Present address: The Rockefeller University, 1230 York Avenue, New York, New York 10021-6399, USA

THE *ras* proto-oncogene products are plasma membrane-bound, guanine nucleotide-binding proteins implicated in signal transduction across the plasma membrane¹. But the signal(s) that activates the *ras* pathway(s) is not known. In the budding yeast *Saccharomyces cerevisiae*, the *CDC25* gene product acts upstream of Ras proteins^{2,3}, but it has not been clear whether *CDC25* function is unique to the *S. cerevisiae* *ras* pathway. Here we report that the *ste6* gene of fission yeast *Schizosaccharomyces pombe* is a homologue of *CDC25*: the *ste6* gene product and the *CDC25* gene product have significant amino-acid similarity in their C-terminal regions. Like the *S. pombe* *ras1* gene^{4,5}, *ste6* is essential for mating⁶. Epistatic interactions indicate that the *ste6* gene functions upstream of *ras1*. We propose that *ste6* and *CDC25* activate Ras protein through a common mechanism, perhaps by promoting GDP-GTP exchange, even though it seems that the function of Ras protein in budding yeast differs from that in fission yeast. Homologues of *ste6* and *CDC25* could regulate *ras* activity in other eukaryotic cells.

Among the sterile (*ste*) genes identified in *S. pombe*, the proto-oncogene homologue *ras1* (allelic with *ste5*; ref. 7) and *ste6* are essential only for mating⁴⁻⁶, whereas the other *ste* genes are also required for meiosis^{6,8,9}. To investigate further the mating process, we isolated the *ste6* gene by screening a *S. pombe* gene library for plasmids able to restore mating to a *ste6*

mutant strain. One such plasmid (pSTE6-1) was obtained, and a 3.27-kilobase (kb) *EcoRI* restriction fragment derived from it rescued a *ste6* mutant strain (Fig. 1). A 2.73-kb open reading frame in this fragment was shown by gene disruption¹⁰ to correspond to the *ste6* gene (Fig. 1). Disruption of the *ste6* gene resulted in the same phenotype as that of the original *ste6* mutant; that is, mating-defective as a haploid but meiosis-competent as a homozygous diploid. The *ste6* gene is predicted to encode a polypeptide of 911 amino acids with a relative molecular mass (M_r) of 105,185 (Fig. 2). The deduced protein sequence was compared¹¹ with sequences in the National Bio-medical Research Foundation (NBRF) database (release 20). The only significant similarity was to the predicted *CDC25* gene product^{2,12} of *S. cerevisiae*. Over 270 amino acids, the amino-acid identity between the C-terminal regions of the deduced products of *ste6* and *CDC25* is 34.4% (Fig. 3). The *SCD25* gene¹³ of *S. cerevisiae*, identified as a multicopy suppressor of a *cdc25* mutant, also has significant similarity to *ste6*; but the two *S. cerevisiae* genes are clearly more closely related to each other than either is to *ste6* (Fig. 3). The N-terminal region of the *ste6* gene product has no obvious similarity with the predicted product of *CDC25*, or with any other protein in the NBRF database. The C-terminal part of *CDC25* is sufficient to rescue a *cdc25* mutant when expressed on a multicopy plasmid^{2,12} indicating that this region has the 'catalytic' activity of the *CDC25*-encoded protein. The homology with the *ste6* gene product is in this region, indicating that the two proteins could have a similar biochemical function.

The similarity between the products of *ste6* and *CDC25* is intriguing, because the *CDC25* gene encodes an activator of the *ras*/adenylyl cyclase pathway in *S. cerevisiae*^{2,3,12}. The *RAS2*^{Val19} mutation, corresponding to the mammalian *ras*^{Val12} oncogene, can suppress the *cdc25* defect^{2,3}, showing that *CDC25* functions upstream of *ras* in *S. cerevisiae*. To determine whether a similar relationship exists between the *S. pombe* *ste6* and *ras1* genes we constructed a double mutant of *ste6* with the *ras1*^{Val17} allele⁵ (corresponding to the mammalian *ras*^{Val12} oncogene). When they were nitrogen-starved, wild-type cells mated to form diploid zygotes (Fig. 4a), whereas the *ste6* mutant strain showed

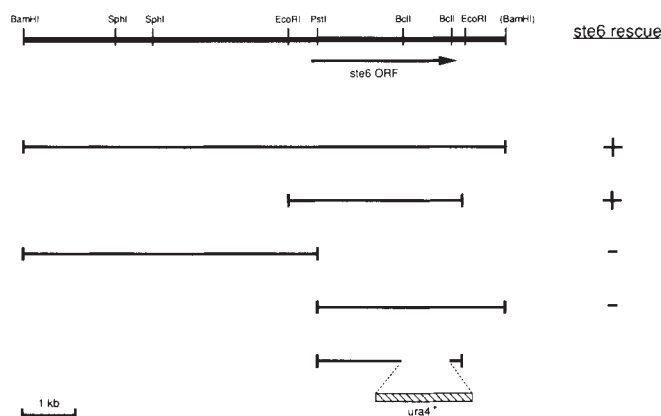


FIG. 1 Molecular cloning and gene disruption of the *ste6* gene. A restriction map of the insert DNA in plasmid pSTE6-1 is shown with the extent and direction of the *ste6* ORF indicated by the arrow (see Fig. 2). Only two sites are shown for the restriction enzymes *BclI* and *EcoRI*. The ability of various subclones to rescue the mating defect of a *ste6* mutant strain is indicated (+ or -). A deletion mutant of *ste6* was constructed *in vivo* by gene replacement. The structure of the linear fragment used to disrupt the *ste6* gene is shown at the foot of the figure.

METHODS. An *h⁹⁰ ste6 ade6 leu1* strain of *S. pombe* was transformed with an *S. pombe* genomic library in the plasmid pDB248' shuttle vector²⁰. *Leu⁺* transformants were tested for mating and sporulation by exposure to iodine vapour, which stains spores dark brown²¹. One transformant with zygotic asci was identified, and this contained the plasmid pSTE6-1. Restriction fragments were subcloned into plasmid pUC118 and tested for rescue of the *ste6* mutant defect by co-transformation with plasmid pDB248'. A 3.27-kb *EcoRI* fragment was sufficient to rescue the *ste6* mutant, and a *PstI* site in this fragment was essential for this activity. The *EcoRI* fragment contained a 2.73-kb open reading frame (ORF) passing through the *PstI* site (Fig. 2). To determine whether this ORF corresponded to the *ste6* gene a one-step gene disruption¹⁰ was performed. Part of the putative *ste6* ORF was deleted by inserting the *ura4⁺* gene²² between two *BclI* restriction sites (note that the deleted part includes the region of homology with the *CDC25* and *SCD25* genes; Fig. 3). A linear *PstI*-*EcoRI* fragment carrying the deleted gene was used to transform²³ an *h⁹⁰ ade6 leu1 ura4* strain. Most of the mitotically stable *Ura⁺* transformants were sterile. Southern-blot analysis of some of these showed that they carried a precise replacement of the wild-type gene by the *ura4⁺*-disrupted gene (data not shown). A disrupted strain was crossed by cell fusion²⁴ to the original *ste6* mutant strain. No *Ste⁺ Ura⁺* progeny were found in 29 tetrads or 350 spores, showing that the disrupted gene is allelic with, or very closely linked to, the *ste6* gene. Media for growth and sporulation, and genetic methods were as previously described^{17,21}.

EcoRI
 -482 GAATTCACATTTTTTTTTTGAAGAAACGCTACCAAGGACAACTCTGTTTTATTTTCAGGAAGAACATTTTCTTCTGATTTATTGAATAGCGAAGGATAGGGTTGAATAGTAAAA
 -362 CAACGAAGTATTACTGTATTGCGGAATAGTCGCATACCCACAGGTTTTAGGTGAGTCCATTGTGAAGTGATTGACACATAGTCTAGTCAATACTGTATGTAAATACAGTTGCGTA
 -242 GTGCAAAACGCGACGCCCAATTCACAGGTAATTTCTATTGTAGTAGCCCATTTGTAGTAAACAAAGAAATTCCTCAAAATGGCTAATATGGGAAAAACAGCTTTTTTTTAAATAAA
 -122 CGGGTGTTCCTTTATTTTAAAGCAGGAAATACGAAATTAAGAAAGGTATATTCCTCTTTTCTTCAAAAATAAATGAATCGTGAATTGATAGCATTCATGGACCCCTTACGA
 PstI
 -2 AAATGAGGTTTCAACGACGCGAATAAGTATTATGAGAACAGCTCTAATCCTTCATTTTAAAAATTTCTGCGAGGACACTATCATAGTTATAGAAGTCTGCGAGGACGGTTGGTGGC
 1 M R F Q T T A I S D Y E N S S N P S F L K F S A G D T I I V I E V L E D G W C D
 122 ACGGAATTTGCTCAGAAAAACGAGGTTGGTTCGCCAGCTGCGTGCATTGATCTTCAAAAATTCAAAATTTTTTTTCAAGTTTTTCATTCATGAATGAGAAAACCCAAATGCTCAATGTT
 41 G I C S E K R G W F P T S C I D S S K I Q N F F S S F H S S N E K D P N A Q C C
 242 GTGCGCGGTTTCAGTAGAGGCTCATCTTCAAGATTCTGCATGGTTTGAGAAAACGCGAGTGAACGATAAATAGTATCCCTTCTCAGAAGAGTTCTTAAGAAAAATCTTCAAAATG
 81 A P F H V E A H L Q D S A W F E K H G V Q A I N S I P S S E E F L R K N L Q N D
 362 ACATTCACCACCTTTGTAAGGAGTTCTCACCACGCTGCGCGTGTGCACAATCTATAAAAGGAAGGCACTCAAGTATGCTTTTTGGAATGAAATGTTCTGATGATGTTCTTT
 121 I H H L V K G I L T T A A A V S Q S I K K E G T Q V I V F G I E T V R S M V L S
 482 CATTTCCCTTGATAATCCTTTCTACATTAGATGAAAATTTTCTCTCAGAAGTGGCGAAGTGTCTCTCATTAATTTATGGCAGAGTTGAGCCGAATGGGTTGCATTATGGTGAAC
 161 F P L I I L S T L D E N F L S E V A Q V F S S I N L L P E L S R M G C T Y G E L
 602 TTTGCATCAGATTACTAAGCTTTTGAAGCAATGGCTAATAAGTTTTGTTTTCTCAGGCCGATGTTTCTTCCCTTCTACTTTTGGGCTCTTGATAGCGCATGAAATACATT
 201 C I R F T K L L K Q L A N K F L F F F R P D V S F P S Y F L G S L I A H E I H F
 722 TCTTGCCATGGGATTTAATATGCTCTGTTCCAATTCTGTACAATCAGCACAATAATCTCAACCTGATATTACTTCTTTGTTGAATTTTGTCACTTTTACACGAAGCTTACCATT
 241 L P W D F N M L C S N S V Q S A H T N L Q P D I T S F V A I L S L S H E A Y H C
 842 GCACTGAGAATGAATTTGGAATTTAGAAGCAGAGAAGCTAACTGAAAATACACCCAAAAAGTACTACAGCTAGTTGCGGAAGATGCACTAGAAGCTTGGAACTAGATATTCTAGAGG
 281 T E N E F W N L E A Q K L T E N T T Q K V L Q L V A E D A L E A W K L D I L E D
 962 ACATCGATAGATGATCAATGTTGTAGCGGATCTTGTCTGCAAAATCAAGAAATAAATATCTTCTCTGAAAATAACCCCTTTTCTTCACTTCTCAAGATGTTGAAGCTTGAAGG
 321 I D R C I Q C C R R F L S A N Q R I N Y S S S E N N P F S F T S Q D V E A L K D
 1082 ATGAAGTCTCTTCAACTATGATGATTTATTTTGTGGAGTATCGACTGGAGCAATCTCACCTAGCGATGTTTACTGGCAATTTTCCCTTTTGTGATTTACTAGTAACCTTGA
 361 E L S S N L C D L Y L W S I D L E Q I S P S D C L L D N Y S L F V D L L V T L K
 1202 AAGTATCCCTTCTCGGATCAAGTCAATAATGTTCAATTTTTCAGAAAGATGTTGTTCTTCTAGATAAATCTCTCACAATATCCAACGAGAATGATGATGCGGAGAAGT
 401 V S L L R I K S I I V Q F S E R I V F L S L E Y K F L T N I Q P E L N D A E K S
 1322 CCCAATCTGATGGTTTGTACCTCAATAAAACCACTGGTTGCACTCTAAAGGATAGTTGTTATTTAATGAACAGACTTCCACGAGGCCATTATGATCGGAACCTTTTGTTCAT
 441 Q L D G F D L N K T N W F D S K G L V C Y L M K Q T S P E P L L I R N L L F S F
 1442 TTGGTCTAGTAATGTAATAAATGAACAAGATGGAATAAATAACAGCCACTTTAGTGTTCATTATAAATACCTTCTAAGGACAGATATAGATAGTACATTTTACTACTATCTTT
 481 W S C N G K I E Q D G K I K T A T L V F I I N Y L L R T D I D S T F F T T I F L
 1562 TAAACACATACGCTAGTATGATGACAGTTCTTTCAGATTATTTTCCACTACTGGAGCACATTTTCGGTTTCATCTGCTCATTAAATTTTGGAAAAATTTCTTTTATTTCTCAGCAATTTTACC
 521 N T Y A S M I S S S D L F S I L G A H F R F I C S L N F G K I S F I S H E F Y R
 1682 GAGTAGTAAGAGGTTTTGGATATACTCTTATTTGGTTGCAATCGTATCTGTTGAAGAGTTGGACAATTCGAAGTCAATTTCTTTTGTAAATTTATAGTTTGAAGTCT
 561 V S K R F L D I L L I W F E S Y L V E E L D N S K S I F F L F K I Y K V F E V F
 1802 TTGATGTTCCACATTTTGATCTGCTGAAGAATTATTCATCTTTTACACCTACTTCATCATCCCTCTACAAAAAGATCACATAAATGCTAGAGGGAAGAGCTATCCCAAGAT
 601 V V P H F A S A E E L L H S L S H L L H H P S T K R S H K M L E G K E L S Q E L
 1922 TAGAGGATCTTTCTCCATAATCCCTGATCAATATATATAAGGATGAATGGTTTACTTCTACCTCTCGTGAAATGCAAGAGCTTATGATCTTAGAGTTTCAATCATTTT
 641 E D L S L I H N S P D P I I Y K D E A L L L P P R I A K Q L I E F Q S F S
 2042 CACACATATCAAGGATTCAGTTCTCTAATAAATCTGGGCAATCTTAACAGATCTCCACCAAGAAAAAATCTGACCTTTTATTTGTCGAATCATCTGGTTAACTTTGTGACCGAAA
 681 H I S R I Q F L T K I W D N L N R F S P K E K T S T F Y L S N H L V N F V T E T
 2162 CCATCGTGAAGAAGAAGAACCTCGCAGAGTACCAATGTGCTAGCATATTTTATCAGGCTGTGATTATTTGAGAGAGCTTAAACAATTTGCTAGTTTATTTCCATCATTTCTGCT
 721 I V Q E E P R R R T N V L A Y F I Q V C D Y L R E L N N F A S L F S I I S A L
 2282 TAAATCTCACCATTTCATCGGCTGCTGAGACATGGGCAATTTGAATAGTAAACATTTGGCTAGTTTGGAGCTTCTAACAATTTGACAGAGCAAGGAAAAATTTTCAGTAATTATA
 761 N S S P I H R L R K T W A N L N S K T L A S F E L L N N L T E A R K N F S N Y R
 2402 GAGATGCTGGAAGAGTGTGCTTGCATGTGCTCTTCTAGGTGTTTACTTCTGATCTGACTTCTCTTAAATCGGAAATAAGATAACTTTCAAAACATGATCAATTTTCGATA
 801 D C I E N C V L P C V P F L G V Y F T D L T F L K T G N K D N F Q N M I N F D K
 2522 AGGCAACCAAGTCACTAGAAATTTGAATGAGATAAAAAAGTTTCAATCTGTTGGGTACATGTTAATCCCATCAACGAAGTTCAAGAGCTTCTTAATGAAGTTATATCGAGAGAGCGAA
 841 R T K V T R I L N E I K F Q S V G Y M F N P I N E V Q E L L N E V I S R E R S
 2642 ACACGAATAACATCTATCAAGAAGTTTAACTGTAGAACCAGTGAATCTGAAGATCAAGCCTTACAACGCTTGCTAATTGATCTGGCATTTTGAAGCGTGAACGTTAACAGTGATT
 881 T N N I Y Q R S L T V E P R E S E D Q A L Q R L L I D S G I F *
 EcoRI
 2762 TAAGTTTTATGAGCTTGCTTCGAATTC

FIG. 2 DNA sequence and corresponding deduced amino-acid sequence (single-letter code) of the *ste6* gene. The DNA sequence of the 3.27-kb *EcoRI* fragment is shown that could encode a 2,733 base ORF passing through the *PstI* site essential for *ste6* function. The deduced amino-acid sequence is shown below the DNA sequence starting with the first methionine codon in the ORF.

METHODS. The DNA sequence was determined on both strands using the dideoxy-chain-termination method²⁵ with modified T7 DNA polymerase²⁶. Sub-clones for sequencing were generated by progressive deletion with exonuclease III/SI nuclease²⁷ from genomic clones in the plasmid pUC118/119 vectors²⁸; gaps in the sequence were filled using oligonucleotide primers complementary to previously determined sequences.

no mating response (Fig. 4b). The *ras1*^{Val17} strain has a characteristic phenotype when nitrogen-starved: cells bearing the *ras1*^{Val17} allele agglutinate more strongly than do wild-type cells, produce long conjugation tubes, and mate poorly (Fig. 4c). This phenotype is dependent on the presence of cells secreting the *M* (minus) mating factor, and indicates that *ras1* regulates the mating factor-response pathway^{4,5,14}. The double-mutant phenotype was similar to that of the *ras1*^{Val17} strain (Fig. 4d). Similar results were obtained using the *ras1*^{Leu66} activated allele (corresponding to the mammalian *ras*^{Leu61} oncogene) (data not shown). The epistatic effect of *ras1*^{Val17} and *ras1*^{Leu66} to the *ste6*

mutation indicates that *ste6* functions upstream of *ras1* in a pathway leading to mating.

The nature of the interaction between *ste6* and *ras1* is unclear. The most pertinent result is that activation of the *Ras* protein by inhibition of its intrinsic GTPase activity (in the *ras1*^{Val17} and *ras1*^{Leu66} mutants) restores the mating-factor response in *ste6* mutants, whereas overexpression of the wild-type *ras1* gene does not (data not shown). This is consistent with the *ste6* gene product functioning as a GDP-GTP exchange factor for the *Ras* protein, as has been suggested for the *CDC25* gene product^{2,3,15,16}. But this interpretation of *ste6* function is

FIG. 3 Amino-acid similarities between the C-terminal regions of the *S. pombe* *ste6* gene product and the *S. cerevisiae* *CDC25* and *SCD25* gene products. Identical amino acids are represented in bold type, and residues conserved in all three gene products are indicated by an asterisk above the *ste6* sequence. The sequences are numbered from the first methionine residue. The alignment between *ste6* and *CDC25* was generated by the FASTP program¹¹, and that between *CDC25* and *SCD25* was from refs 13 and 29.

<i>ste6</i>	640	LEDLSLHNSPDPIIYK.DELVLL.LPPREIAKQCLCILEFQSFSSHISRIQF
<i>CDC25</i>	1279	MSSSSSLPSSASSAFFRLKLLKLLDIDPYTYATQTLVLEHDLYLRLITMPEC
<i>SCD25</i>	947	LAVDPVLFATQLTILEHEIYCEITTFDC
<i>ste6</i>	688	*LTKIWD.NLNR.F.SPKEKTSTFYLSNH.LVNFVTEITVQKEEPRRRRTNVL
<i>CDC25</i>	1329	LDRAWGTKYCNM.GGSPNITKFIAMANTLTNFVSHITVQKADVKTTRSKLT
<i>SCD25</i>	975	LQKIWKNKYTKSYGASPLNEFISFANKLTNFIISYSVKKEADKSKRAKLL
<i>ste6</i>	735	*A YFIQVCDYLRRLNNFASLFSIISALNSSPIHRLRKTWANLNSKTLASFEE
<i>CDC25</i>	1378	QYFVTVAQHCKELNNFSSMTAIVSALYSSPIYRLKKTWDLVSTESKDLK
<i>SCD25</i>	1025	SHFIFIIEYCRKFNNFSSMTDIISALYSSPIYRLKKTWQAVIPQTRDILLQ
<i>ste6</i>	785	*LNNLTLEARKNFPSNYRDCLENCV.LPCVPFLGVYFTDLTFLKTKGNKDNFQ
<i>CDC25</i>	1428	NLNNLMDSKRNPFVKYRELLRSVTDVACVPPFFGVYFTDLTFTFVGNPDLF.
<i>SCD25</i>	1075	SLNKLMDPKKNFINYRNELEKSLHSA PCVPFFGVYLSLDTFTTDSGNPDYLV
<i>ste6</i>	834	NMINFDRKRTKVTRILNEIKKKFQSVGYMFPNPINEVQELL
<i>CDC25</i>	1477	HNSTNIINFSLKRTKIANIVEEIIISPKRFHYKLKRLDDIQTVI
<i>SCD25</i>	1125	LEHGLKGVDHEKKYINFNKRSLVDILQEIYIFKKTHYDFTKDRTVIECI
<i>ste6</i>	872	NEVISRERNNTNNIYQRLTVEPRESEDQALQRLLLID.SGIF
<i>CDC25</i>	1519	EASLENVPHIEKQYQLSLQVEPRSGNTKSGTHASSA.SGTEK
<i>SCD25</i>	1175	SNLENIPIHIEKQYQLSLIIEPKPRKKVVPNSNSNNKSQEK

open to question because *ste6*⁻ mutants are defective only in mating, whereas *ras1*⁻ mutants are also inhibited in meiosis and have a deformed cell shape. The inhibition of sporulation in *ras1*⁻ diploids is likely to be due to an effect on mating-type gene expression (Y. Watanabe and M. Y., unpublished results); the cause of the spherical cell shape in *ras1*⁻ cells is unknown. A low level of *ras1* activity, independent of *ste6* function, could

be sufficient to promote sporulation and normal cell shape. Alternatively, there could be activators of *ras1* specific for meiosis or vegetative growth. The *ral* genes^{17,18} function upstream of *ras1*, but preliminary results indicate that they function independently of *ste6* (our unpublished results). Mating of *S. pombe* is initiated only after nitrogen starvation, and is inhibited in rich media¹⁹. We speculate that the *ste6* gene product activates the *Ras* protein in response to nitrogen deprivation. The unique N-terminal regions of *ste6* and *CDC25* could be responsible for the differences in the regulation of *ras* activity between the two yeasts.

Our results show that although the *ras* genes of *S. pombe* and *S. cerevisiae* have quite different functions, at least one function upstream of *ras*, encoded both by *ste6* and *CDC25*, is conserved. The isolation of a mammalian homologue of *ste6* and *CDC25* should greatly assist the elucidation of the signal transduction pathways that act through the *Ras* proteins. □

Received 13 December 1989; accepted 30 January 1990.

- Barbacid, M. A. *Rev. Biochem.* **56**, 779–827 (1987).
- Broeck, D. et al. *Cell* **48**, 789–799 (1987).
- Robinson, L. C., Gibbs, J. B., Marshall, M. S. & Tatchell, K. *Science* **235**, 1218–1221 (1987).
- Fukui, Y., Kozasa, T., Kaziro, Y., Takeda, T. & Yamamoto, M. *Cell* **44**, 329–336 (1986).
- Nadin-Davis, S. A., Nasim, A. & Beach, D. *EMBO J.* **5**, 2963–2971 (1986).
- Girgsdies, O. *Curr. Genet.* **6**, 223–227 (1982).
- Lund, P. M. et al. *Molec. gen. Genet.* **209**, 627–629 (1987).
- Thuriaux, P., Sipiczki, M. & Fantès, P. A. *J. gen. Microbiol.* **116**, 525–528 (1980).
- Michael, H. & Gutz, H. *Yeast* **3**, 5–9 (1987).
- Rothstein, R. *Meth. Enzym.* **101**, 202–211 (1983).
- Lipman, W. J. & Pearson, W. R. *Science* **227**, 1435–1441 (1985).
- Camonis, J. H. et al. *EMBO J.* **5**, 375–380 (1986).
- Boy-Marcotte, E., Damak, F., Camonis, J., Garreau, H. & Jacquet, M. *Gene* **77**, 21–30 (1989).
- Fukui, Y., Kaziro, Y. & Yamamoto, M. *EMBO J.* **5**, 1991–1993 (1986).
- Camonis, J. & Jacquet, M. *Molec. cell. Biol.* **8**, 2980–2983 (1986).
- Powers, S., O'Neill, K. & Wigler, M. *Molec. cell. Biol.* **9**, 390–395 (1989).
- Fukui, Y. & Yamamoto, M. *Molec. gen. Genet.* **215**, 26–31 (1988).
- Fukui, Y., Miyake, S., Satoh, M. & Yamamoto, M. *Molec. cell. Biol.* **9**, 5617–5622 (1989).
- Egel, R. *Planta* **98**, 89–96 (1971).
- Beach, D., Piper, M. & Nurse, P. *Molec. gen. Genet.* **187**, 326–329 (1982).
- Gutz, H., Heslot, H., Leupold, U. & Lopreano, N. in *Handbook of Genetics* Vol. 1 (ed. King, R. D.) 395–446 (Plenum, New York, 1974).
- Bach, M.-L. *Curr. Genet.* **12**, 527–534 (1987).
- Ito, H., Fukada, Y., Murata, K. & Kimura, A. *J. Bact.* **153**, 163–168 (1983).
- Sipiczki, M. & Ferenczy, L. *Molec. gen. Genet.* **151**, 77–81 (1977).
- Sanger, F., Nicklen, S. & Coulson, A. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
- Tabor, S. & Richardson, C. C. *Proc. natn. Acad. Sci. U.S.A.* **84**, 4767–4771 (1987).
- Henikoff, S. *Gene* **28**, 351–359 (1983).
- Vieira, J. & Messing, J. *Meth. Enzym.* **153**, 3–11 (1987).
- Jacquet, M., Camonis, J., Boy-Marcotte, E., Damak, F. & Garreau, H. in *The Guanine-Nucleotide Binding Proteins* (eds. Bosch, L., Kraal, B. & Parmeggiani, A.) 241–249 (Plenum, New York, 1989).

ACKNOWLEDGEMENTS. We thank Drs C. Shimoda and S. Nadin-Davis for fission yeast strains. D.A.H. was supported by the Royal Society and the Japan Society for the Promotion of Science. This work was supported by the Ministry of Education, Science and Culture of Japan.

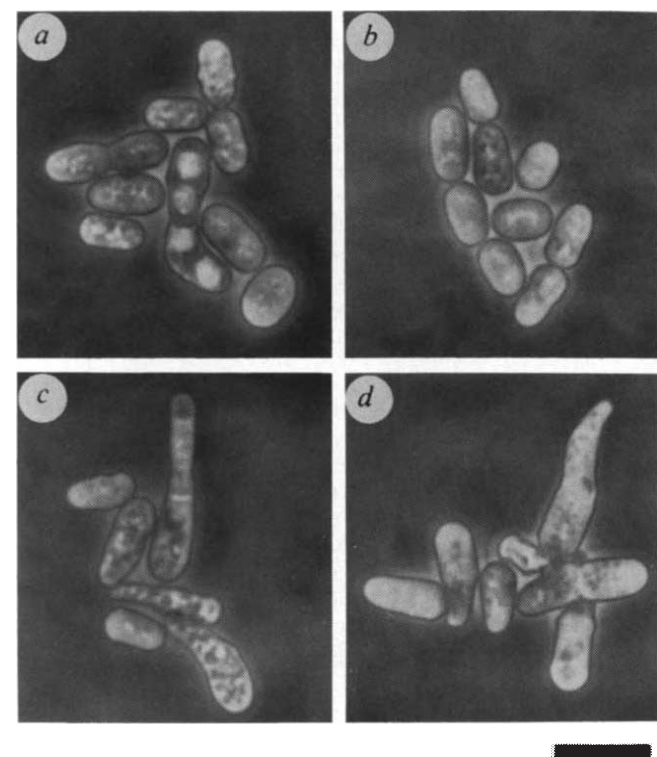


FIG. 4 Epistatic effect of the *ras1*^{Val17} allele on the *ste6* mutation. Phase-contrast micrographs of wild-type (*h*⁹⁰ *ade6 leu1*) cells (a); *ste6* disruptant (*h*⁹⁰ *ste6::ura4*⁺ *ade6 leu1 ura4*) cells (b); *ras1*^{Val17} mutant (*h*⁹⁰ *ras1*^{Val17} < *LEU2 ade6 leu1*) cells (c); and *ste6 ras1*^{Val17} double-mutant (*h*⁹⁰ *ste6::ura4*⁺ *ras1*^{Val17} < *LEU2 ade6 leu1 ura4*) cells (d); Scale bar, 10 μ m. METHODS. Double mutants were constructed by cell fusion²⁴ and tetrad analysis²¹. Cell cultures were grown to stationary phase in nitrogen-limited liquid medium¹⁹ at 28 °C.

Sequence-specific artificial photo-induced endonucleases based on triple helix-forming oligonucleotides

L. Perrouault*, U. Asseline†, Christian Rivalle‡, N. T. Thuong†, E. Bisagni‡, C. Giovannangeli*, T. Le Doan* & C. Hélène*

* Laboratoire de Biophysique, Muséum National d'Histoire Naturelle, INSERM U.201—CNRS UA.481, 43 rue Cuvier, 75005 Paris, France
† Centre de Biophysique Moléculaire, CNRS, 45071 Orléans Cedex 02, France

‡ Institut Curie, URA 1387 CNRS, Bâtiment 110, 91405 Orsay, France

HOMOPYRIMIDINE oligonucleotides bind to homopurine-homopyrimidine sequences of duplex DNA forming a local triple helix^{1–8}. This binding can be demonstrated either directly by a footprinting technique³, gel assays⁴, or indirectly by inducing irreversible reactions in the target sequence, such as photocrosslinking^{1,5} or cleavage^{2,6–8}. Binding occurs in the major groove with the homopyrimidine oligonucleotide orientated parallel to the homopurine strand. Thymine and protonated cytosine in the oligonucleotide form Hoogsteen-type hydrogen bonds with A·T and G·C Watson–Crick base pairs, respectively. Here we report that an 11-residue homopyrimidine oligonucleotide covalently attached to an ellipticine derivative by its 3' phosphate photo-induces cleavage of the two strands of a target homopurine-homopyrimidine sequence. To our knowledge, this is the first reported case of a sequence-specific artificial photoendonuclease. In addition we show that a strong binding site for a free ellipticine derivative is induced at the junction between the triplex and duplex structures on the 5' side of the bound oligonucleotide. On irradiation, cleavage is observed on both strands of DNA. This opens new possibilities for inducing irreversible reactions on DNA at specific sites by the synergistic action of a triple helix-forming oligonucleotide and an intercalating agent.

The attachment of photoactive groups to oligonucleotides can induce site-specific photocrosslinking reactions of the oligonucleotide to its target^{1,5,9,10}. The photocrosslinks produced by azido derivatives, azidoproflavine¹ or azidophenacyl⁵, can be converted to strand breaks under alkaline conditions, but none of the photosensitizers induce direct cleavage of the target nucleic acid. We attached an ellipticine derivative (Fig. 1) to the 3' end of an 11-nucleotide (11-mer) homopyrimidine oligonucleotide 5'-d(TTCCTCCTCT)-3'. This oligonucleotide binds to a homopurine-homopyrimidine sequence, 11 base pairs in length, contained in a 32-bp DNA fragment (Fig. 2). On irradiation at wavelengths >300 nm, we observed photoinduced cleavage of both strands of duplex DNA (Figs 2 and 3) at the position expected if the oligonucleotide binds in a parallel orientation with respect to the homopurine-containing strand. A small fraction of the oligonucleotide was photocrosslinked to the duplex DNA target, as revealed by species migrating more

slowly than the original fragment on neutral gels (Fig. 3). Temperature studies showed that both photocrosslinking and cleavage disappeared when the oligonucleotide–ellipticine conjugate was dissociated from the 32-bp DNA fragment (50% dissociation was observed at 22 °C).

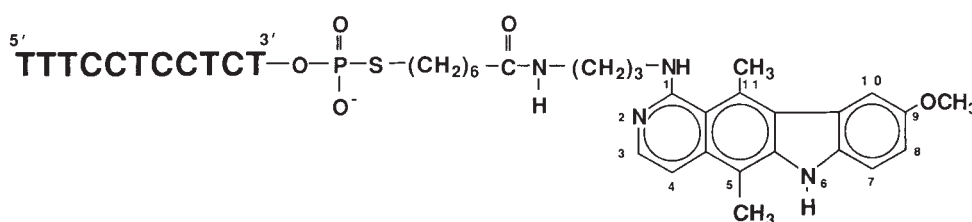
Irradiation of the 32-bp duplex in the presence of 5 µM of the free ellipticine derivative used to synthesize the 11-mer-ellipticine conjugate (with a –NH–(CH₂)₃–NH₂ substituent in position 1; see Fig. 1) resulted in cleavage of both strands at most positions, with a preference for guanines (Fig. 4). This could reflect the slight preference of ellipticine derivatives to intercalate at G·C base pairs¹¹. When we first incubated the 32-bp duplex with the unsubstituted 11-mer homopyrimidine oligonucleotide, strong cleavage was induced by free ellipticine derivative on both sides of the bound 11-mer oligonucleotide (Fig. 4). The strongest cleavage was observed at the TpA sequence that is present on both strands of the 32-bp DNA fragment on the 5' side of the 11-mer oligonucleotide at the triplex–duplex junction. This cleavage was observed even at low ellipticine concentrations (0.2 µM), at which there was almost no cleavage of the 32-bp duplex in the absence of the 11-mer oligonucleotide (Fig. 4).

Binding of the 11-mer oligopyrimidine to the major groove of the 32-bp DNA fragment did not prevent cleavage by ellipticine within the triple helical region strongly indicating that the photochemical reaction occurred from the minor groove (Fig. 4). When ellipticine is covalently attached to the homopyrimidine oligonucleotide, it is brought into the major groove by oligonucleotide binding. Cleavage can take place in the minor groove if ellipticine intercalates at the triplex–duplex junction. Model-building studies showed that the linker used to tether ellipticine to the oligonucleotide allowed for intercalation at both the ApT and TpG steps at the triplex–duplex junction on the 3' side of the bound oligonucleotide. An acridine derivative attached to the 5' end of the 11-mer oligonucleotide was previously shown to intercalate at the triplex–duplex junction on the 5' side¹².

The gel mobility of the cleaved fragments showed that two termini were generated (see Fig. 4 for free ellipticine). One of them migrated as phosphomonoesters produced in the Maxam–Gilbert sequencing reaction, whereas the second migrated faster, as did phosphoglycolates produced, for example, by ⁶⁰Co irradiation or copper–phenanthroline cleavage¹³. These two products arise from oxidative attack at C1' and C4' of deoxyribose in the minor groove. The relative amounts of the two products depended on the particular cleavage site. They had nearly equal intensity at the adenosine site at the triplex–duplex junction, whereas there was a large excess of the phosphomonoester at the adjacent thymidine (Fig. 4). The narrow localization of the cleavage sites at the T and A nucleotides on each strand at the triplex–duplex junction (Fig. 4) argues against diffusible species (such as OH radicals) as intermediates in the photochemical reaction. Direct reaction of photoexcited ellipticine with nearest neighbour nucleotides seems, therefore, more likely.

The results presented above show that it is possible to design a site-specific artificial photoendonuclease by attaching an ellipticine derivative to a homopyrimidine oligonucleotide that

FIG. 1 Structure of the ellipticine derivative used to cleave DNA under photochemical activation. 1-Chloro 5,11-dimethyl 9-methoxyellipticine was first reacted with diaminopropane as described previously¹⁴. This derivative was reacted with ω-bromohexanoic acid and the product attached to the 3' thiophosphate group carried by the oligonucleotide, as previously reported¹⁵.



Control of the cleavage reaction is, therefore, much more easily achieved than it is with, for example, metal chelates, which have been previously used to induce site-specific cleavage of double helical DNA when tethered to oligonucleotides^{2,6-8}. We have also shown that local triple helix formation creates a strong

FIG. 2. Sequence of the 32-bp DNA fragment used as a target for the 11-mer oligonucleotide substituted by the ellipticine derivative as shown in Fig. 1. The target homopurine-homopyrimidine sequence is boxed. The bars indicate the efficiency of cleavage observed on irradiation of the triple helix at wavelengths >300 nm. The values were obtained from densitometric analysis of autoradiographs of gels similar to that presented in Fig. 3.

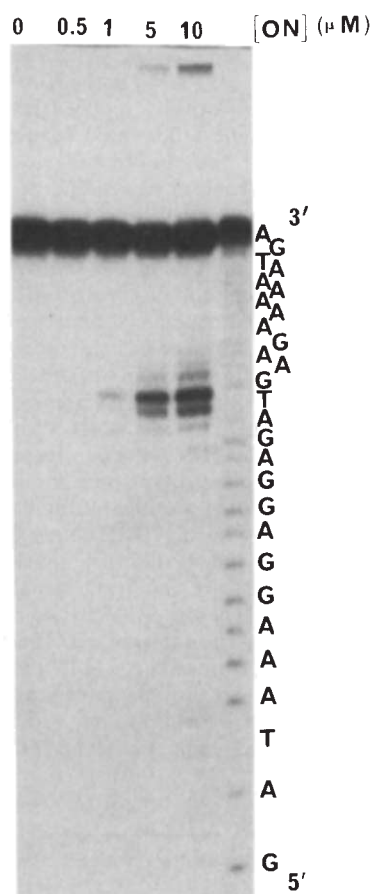
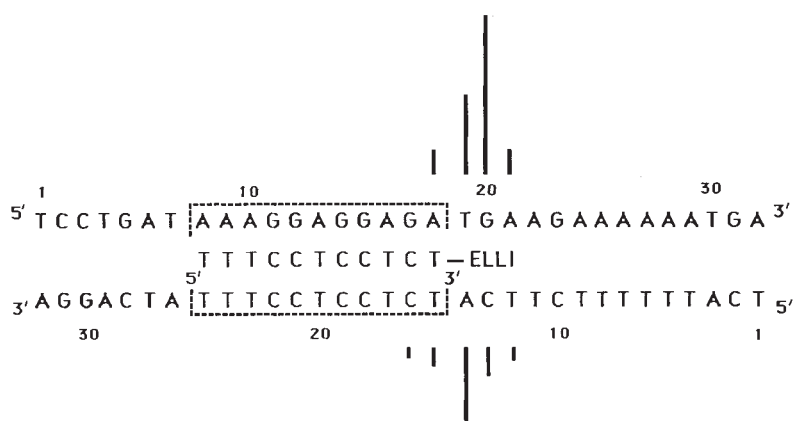


FIG. 3 Photo-induced cleavage of the homopurine-containing strand of the 32-mer DNA fragment irradiated in the presence of the 11-mer ellipticine conjugate. The 32-mer fragment (10 nM) was incubated with increasing concentration of the oligonucleotide-ellipticine conjugate ([ON]) from 0 μ M to 10 μ M (as indicated at the top of the figure) in a pH 6 buffer containing 10 μ M sodium cacodylate and 0.1 M NaCl at 0 $^{\circ}$ C. The mixture was then irradiated for 20 min at wavelengths >300 nm. The irradiated samples were loaded on a denaturing 20% polyacrylamide-7 M urea gel. The gels were autoradiographed with Fuji X-ray films at -20° C. The right lane is the G + A sequencing reaction. Cleavage occurs predominantly at residue G20 and T19 (see Fig. 1). Photo-induced crosslinks are revealed as bands migrating more slowly than the 32-mer (top).

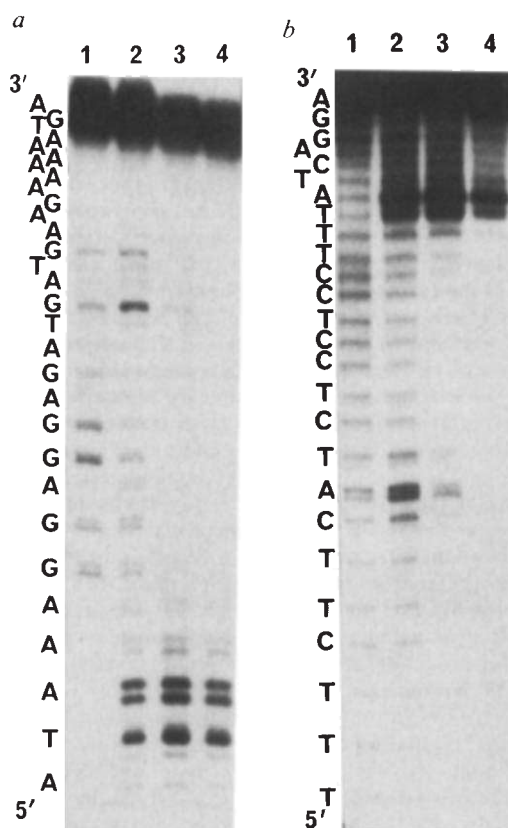


FIG. 4. Triple helix-directed selective cleavage of the 32-mer duplex (Fig. 2). Left, homopurine-containing strand; right, homopyrimidine-containing strand. The 32-mer duplex (10 nM) was incubated with 5 μ M 11-mer unsubstituted oligopyrimidine at 0°C in the buffer described in Fig. 3. The mixture was then irradiated in the presence of 5 μ M (lane 2), 1 μ M (lane 3) and 0.2 μ M (lane 4) ellipticine derivative used to synthesize the derivatized oligonucleotide described in Fig. 1 (the substituent in position 1 of the ellipticine ring was $H_2N-(CH_2)_3-NH-$). Lane 1 is the control irradiated in the absence of 11-mer oligonucleotide and in the presence of 5 μ M ellipticine derivative.

binding site for a free ellipticine derivative at the triplex-duplex junction. This synergistic effect targets the binding of a non-specific ligand to a specific site and points to a new way of eliciting sequence-specific biological effects. □

Received 4 October 1989; accepted 23 January 1990.

1. Le Doan, T. et al. *Nucleic Acids Res.* **15**, 7749–7760 (1987).
2. Moser, H. E. & Dervan, P. B. *Science* **238**, 645–650 (1987).
3. François, J. C., Saison-Behmoaras, T. & Hélène, C. *Nucleic Acids Res.* **16**, 11431–11440 (1988).
4. Lyamyshev, V. I., Mirkin, S. M., Frank-Kamenetskii, D. & Cantor, C. R. *Nucleic Acids Res.* **16**, 2165–2178 (1988).
5. Praseuth, D. et al. *Proc. natn. Acad. Sci. U.S.A.* **85**, 1349–1353 (1988).
6. Povsic, T. J. & Dervan, P. B. *J. Am. chem. Soc.* **111**, 3059–3061 (1988).
7. François, J. C., Saison-Behmoaras, T., Chassinol, M., Thuong, N. T. & Hélène, C. *J. biol. Chem.* **264**, 5891–5898 (1989).
8. François, J. C. et al. *Proc. natn. Acad. Sci. U.S.A.* **86**, 9702–9706 (1989).
9. Lee, B. K., Murakami, A., Blake, K. R., Lin, S. B. & Miller, P. S. *Biochemistry* **27**, 3197–3203 (1988).
10. Praseuth, D. et al. *Biochemistry* **27**, 3031–3038 (1988).
11. Schwaller, M. A., Aubard, J., Auclair, C., Paoletti, C. & Dodin, G. *Eur. J. Biochem.* **181**, 129–134 (1989).
12. Sun, J. S. et al. *Proc. natn. Acad. Sci. U.S.A.* **86**, 9198–9202 (1989).
13. Sigman, D. S. *Accs chem. Res.* **19**, 180–186 (1986).
14. Ducrocq, C. et al. *J. med. Chem.* **23**, 1212–1216 (1980).
15. Asseline, U. & Thuong, N. T. *Tetrahedron Lett.* **30**, 2521–2524 (1989).

ACKNOWLEDGEMENTS. This work was supported in part by Rhône-Loulenc-Santé and the Ligue Nationale Française contre le Cancer.

Progesterone receptor stimulates transcription of mouse mammary tumour virus in a cell-free system

Martha Kalff, Bernhard Gross & Miguel Beato*

Institut für Molekularbiologie und Tumorforschung,
E. Mannkopf-Strasse 2, D-3550 Marburg, FRG

IN higher eukaryotes, steroid receptors are general modulators of gene activity which bind to DNA hormone response elements (HREs) in the vicinity of regulated promoters¹. Analysis of mutant and chimaeric receptor proteins in gene transfer experiments has identified the domains responsible for hormone binding, interaction with the HREs and transactivation², but the transactivating function of hormone receptors has proved difficult to reproduce in cell-free assays. Here we describe a crude *in vitro* system in which transcription from the mouse mammary tumour virus (MMTV) promoter is increased up to 10-fold by native progesterone receptor from rabbit uterus. The stimulatory effect must depend on receptor binding to the MMTV-HRE as it is abolished by deletion or mutation of the HRE and by HRE-oligonucleotide competition. The nuclear factor-I binding site immediately downstream of the HRE, however, does not appear to be essential to progesterone receptor-mediated *in vitro* stimulation of MMTV transcription. The transactivation activity of the progesterone receptor depends on binding of a functional ligand and so this assay should be useful for dissection of the mechanism of transcription activation by steroid hormones.

The MMTV promoter is a well characterized target of transcriptional regulation by steroid hormones *in vivo*¹. To develop a functional assay for transcriptionally active receptors *in vitro*, we studied the transcription of negatively supercoiled MMTV promoter templates in HeLa cell nuclear extracts³. In template-mixing experiments, we find that both the wild-type MMTV promoter containing the intact HRE (+HRE) and a deleted promoter without the HRE (–HRE) are accurately transcribed at low template concentrations when assessed by RNase mapping (Fig. 1a and b, lane 5). Deletion of the HRE sequences without disruption of the nuclear factor-I (NF-I) binding site does not reduce the level of transcription. Therefore, the basal

promoter, although it is almost silent *in vivo*⁴, shows constitutive activity *in vitro*.

We used this sensitive assay to test the efficiency of the MMTV promoter when we added progesterone receptor *in vitro*. The receptor was prepared from rabbit uterus^{5,6} and runs on silver-stained polyacrylamide gels (Fig. 1c) as a single band, with apparent molecular weight 110,000 (110K), which is the intact form⁷; two minor bands probably represent degradation products. All our receptor preparations yield DNase I footprints over the MMTV-HRE⁶. Incubation of the wild-type and –HRE templates with purified progesterone receptor stimulates correct transcription only from the wild-type promoter in a way that depends on receptor concentration (Fig. 1b, lanes 5–8): there is a 10-fold increase in the amount of wild-type MMTV transcript (average of 6 experiments) at the highest receptor concentration. This corresponds at most to an equimolar ratio of receptor to HRE, assuming that a receptor dimer occupies each of the four TGTCT motifs of the HRE⁶.

To ensure that this stimulation was due to receptor binding to the HRE, we included competitor DNAs in the reaction. A competitor consisting of 10 linked copies of the distal MMTV progesterone receptor-binding site⁸ selectively prevents stimulation of the wild-type promoter (Fig. 2a, lanes 5–8), but a non-specific competitor does not (Fig. 2a, lanes 1–4 and 9–12). Furthermore, a mutant template in which the TGTCT motif at nucleotide positions –98 to –93 in the MMTV promoter has been altered shows a drastically reduced response to added receptor, even though the other three hexanucleotide motifs remain intact (Fig. 2b). The *in vitro* behaviour of this mutant is reminiscent of its behaviour in gene transfer experiments^{6,9} and indicates that the functional cooperativity between binding sites *in vivo* is probably preserved *in vitro*.

We next investigated the role of NF-I recognition sequences in MMTV transcriptional enhancement. Binding of this transcription factor at a site immediately downstream of the HRE¹⁰ has been implicated in glucocorticoid induction of MMTV expression^{11,12}. Competition with an oligonucleotide containing the consensus NF-I binding site¹³ decreases activity of both the wild-type and –HRE MMTV promoters to the same extent *in vitro* (data not shown), but it does not interfere with the receptor-dependent stimulation of wild-type transcription (Fig. 3a). We still find a response to progesterone receptor when it is added to a mutant template with the HRE intact but containing 6 substituted base pairs in the NF-I binding site to prevent binding of the factor⁹ (Fig. 3b). It is significant that the amplitudes of the mutant and wild-type responses are comparable, despite reduced basal transcription of the mutant. Thus, although NF-I acts as a transcription factor on the MMTV promoter, the effect of the progesterone receptor is independent of NF-I binding to its cognate sequence in the MMTV long terminal repeat. This is consistent with earlier gene transfer experiments, in which

TABLE 1 Ligand effect on transactivation

Preincubation	Ligand	Fold-induction
0 min	None	8.2 (6.5, 10)
40 min	None	1.3 (1.1, 1.5)
40 min	2.5 × 10 ^{–7} M R5020	6.0 (5.0, 7.0)
40 min	2.5 × 10 ^{–7} M RU486	1.7 (1.6, 1.8)

Freshly thawed aliquots of progesterone receptor were preincubated at 4 °C, before adding template at the times indicated, either with or without R5020 (17,21-dimethyl-19-norpregna-4,9-dien-3,20-dione) or RU486 (17β-hydroxy-11β-(4-dimethylaminophenyl)-17α-(1-propynyl)-oestra-4,9-dien-3-one) present. After 20 min incubation with receptor (15 ng), test (+HRE) and control (–HRE) templates (25 ng each) were transcribed and RNase-mapped as described in the legend to Fig. 1. The induction ratio of control to test template represents the mean of two independent experiments (in parentheses).

* To whom correspondence should be addressed.

mutation of the NF-I recognition sequence of the MMTV promoter was found to inhibit glucocorticoid but not progesterone induction⁹.

We then established the influence of hormone on transcriptional enhancement. We isolated purified progesterone receptor and froze it in the hormone-bound state, but the hormone rapidly dissociates from the receptor on thawing. Dextran-coated char-

coal treatment reveals that after a 40-min incubation, only 7.5% of the receptor is occupied by ligand. This dissociation is partially reversible in that prolonged incubation in the presence of a ³H-labelled synthetic agonist, Org2058, causes reloading of the receptor (data not shown). Addition of Org2058 or another synthetic agonist R5020, to the receptor preparation preserves its transactivation capacity during a 40-min preincubation, but

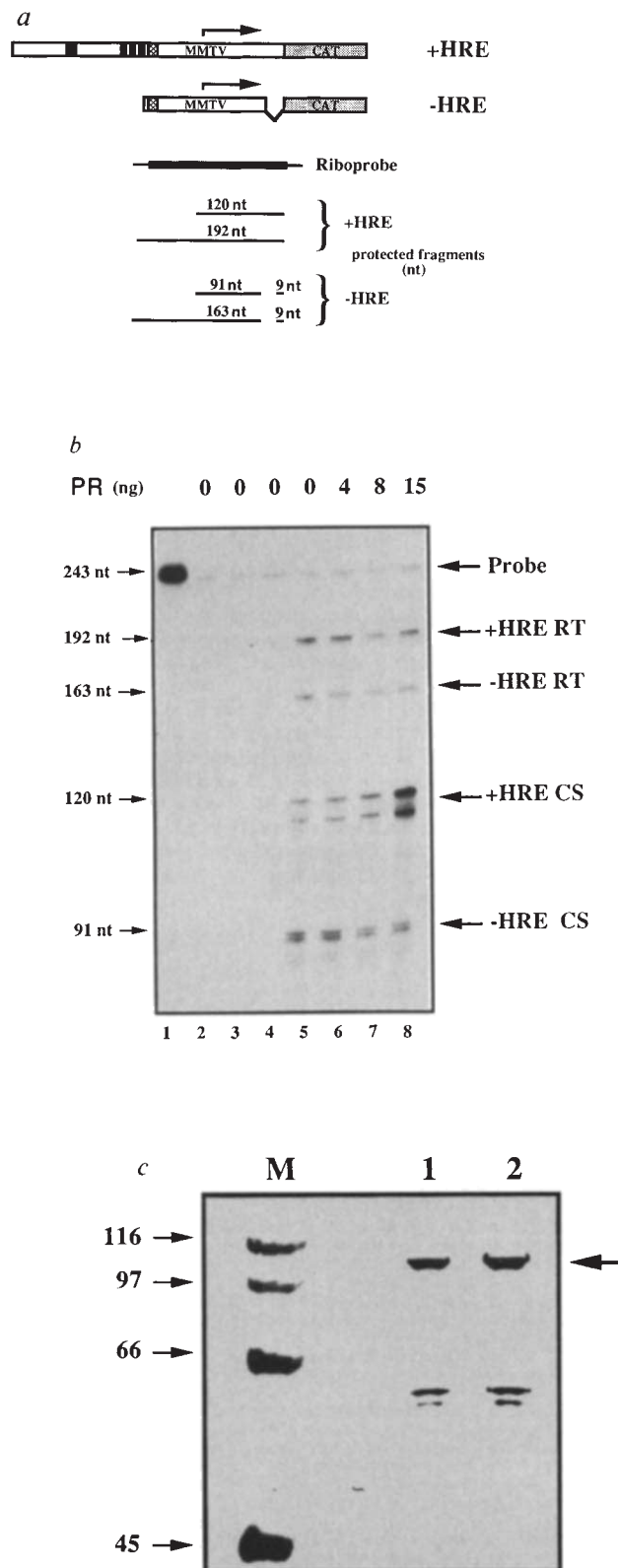


FIG. 1 Influence of purified progesterone receptor (PR) on the activity of the MMTV promoter *in vitro*. **a**, Schematic representation of MMTV promoter templates and RNase mapping. The wild-type (+HRE) plasmid contains MMTV-LTR sequences from -231 to +111 base pairs linked to the chloramphenicol acetyl transferase (CAT) gene. In the -HRE plasmid, MMTV sequences extend from -77 to +91 bp relative to the cap site. Black boxes indicate the progesterone receptor binding sites in the MMTV HRE; stippled boxes, the NF-I binding site. The extent of the single-stranded antisense RNA probe is indicated below: complementary sequences are shown as a thick solid bar and unrelated plasmid polylinker sequences as a thin line. Also shown are the RNase-resistant hybrids predicted to form between the probe and transcripts derived from the wild-type and -HRE templates. Protected fragments from transcripts originating at the correct start site (120 and 91 nucleotides (nt)), as well as read-through (192 and 163 nt) transcripts initiating upstream of the region complementary to the probe, are indicated. Note that the wild-type protected fragments are 9 nucleotides longer than the transcribed MMTV sequences owing to additional complementarity of the probe and CAT sequences. **b**, Quantitative RNase mapping of transcripts synthesized *in vitro* from MMTV plasmid templates containing (+HRE) or missing (-HRE) the HRE sequences. CS, Transcript originates from the correct start site; RT, read-through transcript initiated upstream of the region complementary to the probe. The correct transcription signals appear as doublets. Undigested probe, (lane 1); control reaction without template DNA, (lane 2); control reaction in the absence of ribonucleoside triphosphates (lane 3); transcription in the presence of α -amanitin ($2 \mu\text{g ml}^{-1}$) (lane 4); transcription in the absence (lane 5) or presence of 4, 8 and 15 ng progesterone receptor (lanes 6-8). **c**, Silver-stained SDS-polyacrylamide gel of progesterone receptor after purification from rabbit uterus (lanes 1 and 2, arrowed). Molecular weight markers are indicated in thousands (lane M).

METHODS. HeLa cell nuclear extracts were prepared as described³. Progesterone receptor was isolated from uterine cytosol of rabbits that received daily injections of oestrogen benzoate in the week before being killed. Cytosol preparation and receptor purification have been described^{5,6}. Final purity of our preparations was 80%. Before transcription test and control templates (25 ng of each) were mixed with sheared calf thymus DNA (50 ng) and incubated with increasing amounts (0-15 ng) of progesterone receptor at room temperature for 20 min. The standard *in vitro* transcription reaction (final volume, 25 μl) contained 5 μl nuclear extract (50 μg protein), 50 ng of negatively supercoiled DNA template and 250 ng calf thymus DNA in 22 mM HEPES buffer, pH 7.9, with 7.5% glycerol, 40 mM NaCl, 20 mM KCl, 0.45 mM EDTA, 0.8 mM dithiothreitol, 40 $\mu\text{g ml}^{-1}$ bovine serum albumin, 5 mM MgCl_2 and 400 μM of each of the ribonucleoside triphosphates with the indicated amounts of receptor. After incubation at 30 °C for 45 min, reactions were stopped by addition of 200 μl 0.1 M sodium acetate, 10 mM EDTA, 0.1% SDS and 20 $\mu\text{g ml}^{-1}$ yeast transfer RNA, then extracted with phenol/ CHCl_3 and ethanol-precipitated. Nucleic acids were dissolved in 40 mM Tris-HCl, pH 7.9, 10 mM NaCl, 10 mM MgCl_2 , and treated with 1.5 units of DNase I (RQ1 DNase; Promega) at 37 °C for 30 min. After extraction with phenol/ CHCl_3 and precipitation with ethanol, RNA was hybridized to a [³²P]-labelled MMTV-specific antisense RNA probe. The 243-nucleotide probe contains sequences from -72 to +111 of the MMTV-LTR. RNA probe synthesis and RNase mapping have been described¹⁶. RNase-resistant products were fractionated on 6% denaturing polyacrylamide gels, visualized by autoradiography and quantitated by densitometry.

FIG. 2 Specificity of the progesterone receptor (PR)-mediated *in vitro* stimulation of the MMTV promoter. MMTV transcripts were quantified by RNase mapping as described in the legend to Fig. 1. Doublet bands are marked at the side of the autoradiograph. *a*, Transcriptions of equimolar amounts of wild-type (+HRE) and control (-HRE) templates preincubated with increasing amounts of receptor in the presence of non-specific or specific competitor DNA. Nonspecific competitor: calf thymus DNA (lanes 1-4) or poly(dI.dC) (lanes 9-12). Specific competitor: a 370-bp fragment consisting of 10 linked copies of an oligonucleotide corresponding to the MMTV promoter's distal progesterone receptor-binding site (lanes 5-8). Competition was against 50 ng DNA, corresponding to a 50-fold molar excess over template DNA. *b*, Comparison of the stimulation of the wild-type (+HRE) and a mutant MMTV promoter template by progesterone receptor. Equimolar amounts of wild-type template and a mutant (M-93/-98) in which the hexanucleotide motif TGTTCT between -98 and -93 has been substituted by a *Sma*I recognition sequence⁶, were incubated with the indicated amounts of receptor and then transcribed. A short 29-base pair 3' deletion in the transcribed MMTV sequences of the mutant plasmid enables the correctly initiated wild-type (120-nt protected fragment) and mutant (91-nt protected fragment) transcripts to be distinguished.

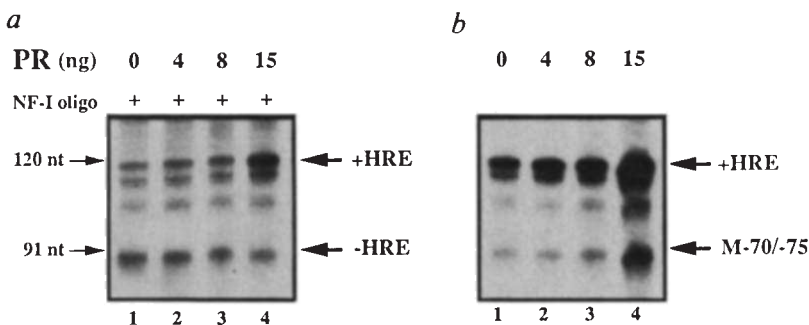
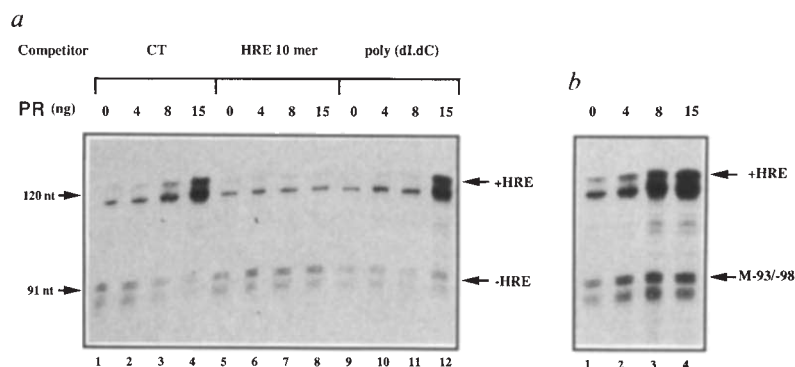


FIG. 3 Influence of the NF-I binding site on the progesterone receptor-stimulated MMTV promoter activity. *a*, Equimolar amounts of wild-type (+HRE) and control (-HRE) template were mixed and incubated with progesterone receptor (PR) in the presence of a 200-fold molar excess of a 35-base pair oligonucleotide (CCTTTTITGGATTGAAGCCAA-TATGATAATGAGG) that contains the NF-I consensus sequence¹³, then transcribed and analysed by RNase mapping as described in the legend to Fig. 1. Positions of the correctly initiated transcripts are indicated. *b*, Wild-type (+HRE) and mutant (M-70/-75) MMTV promoter templates were mixed at equimolar ratios in the presence of calf thymus DNA and the indicated amounts of receptor and then transcribed. The mutant template contains an oligonucleotide substitution between -70 and -75 in the NF-I binding site⁹ and a short 3' deletion in the transcribed MMTV sequences. MMTV transcripts give rise to RNase-resistance bands of 120 nucleotides (+HRE) and 91 nucleotides (M-70/-75), as indicated by the arrows.

this is abolished in the absence of ligand or in presence of the antagonist RU486 (Table 1). Thus, receptor devoid of ligand can bind specifically to DNA⁸, but agonist-loading is necessary for the receptor to act as a transcription factor *in vitro*. The anti-hormone RU486, which is thought to act at a distal stage in progesterone receptor action¹⁴, cannot fulfil this function, so it might be interfering with the transactivation process itself.

In summary, we have used a cell-free transcription assay to test the activity of a purified steroid receptor. The stimulation exerted by the progesterone receptor *in vitro* mimics progesterone induction of MMTV expression *in vivo* in that the HRE must be *in cis*, it depends on binding of an agonist, and it is independent of the transcription factor NF-I. This stimulation is modest in our cell-free system compared with the hormone response *in vivo*, but this can probably be accounted for by our non-saturating receptor concentrations and basal activity of the MMTV promoter, which is silent in the absence of hormone *in vivo*. Thus, any negative regulation that operates in intact cells and involves changes in chromatin structure¹⁵ is not evident in our HeLa cell nuclear extract. It should be feasible to study the role of multiple transactivation domains and the steps in the

hormonal activation of transcription by using mutant receptors as well as templates packaged into chromatin. □

Received 5 October 1989; accepted 5 February 1990.

1. Beato, M. *Cell* **56**, 335-344 (1989).
2. Evans, R. M. *Science* **240**, 889-895 (1988).
3. Dignam, J. D., Lebovitz, R. M. & Roeder, R. G. *Nucleic Acids Res.* **11**, 1475-1489 (1983).
4. Ringold, G. M. *Biochim. biophys. Acta* **560**, 487-508 (1979).
5. von der Ahe, D. *et al. Nature* **313**, 706-709 (1985).
6. Chalepakos, G. *et al. Cell* **53**, 371-382 (1988).
7. Loosfelt, H., Logeat, F., Hai, M. T. V. & Milgrom, E. *J. Biol. Chem.* **259**, 14196-14202 (1984).
8. Schauer, M., Chalepakos, G., Willmann, T. & Beato, M. *Proc. natn. Acad. Sci. U.S.A.* **86**, 1123-1127 (1989).
9. Cato, A. B. C., Skroch, P., Weinmann, J., Butkeraits, P. & Ponta, H. *EMBO J.* **7**, 1403-1410 (1988).
10. Nowock, J., Borgmeyer, U., Püschel, A. W., Rupp, R. A. W. & Sippel, A. E. *Nucleic Acids Res.* **13**, 2045-2061 (1985).
11. Buetti, E. & Kühnel, B. *J. molec. Biol.* **190**, 379-389 (1986).
12. Miksicek, R., Borgmeyer, U. & Nowock, J. *EMBO J.* **6**, 1355-1360 (1987).
13. Meisterernst, M., Gander, I., Rogge, L. & Winnacker, E.-L. *Nucleic Acids Res.* **16**, 4419-4435 (1988).
14. Guiochon-Mantel, A. *et al. Nature* **336**, 695-698 (1988).
15. Piña, B., Brüggemeier, U. & Beato, M. *Cell* (in the press).
16. Zinn, K., DiMaio, D. & Maniatis, T. *Cell* **34**, 865-879 (1983).

ACKNOWLEDGEMENTS. We thank A. Misseyanni for discussion, and R. Haché, E. Slater and G. Suske for critically reading the manuscript. This work was supported by grants from the Deutsche Forschungsgemeinschaft and the Fonds der chemischen Industrie. M.K. holds a long-term EMBO Fellowship.

Reverse hydrophobic effects relieved by amino-acid substitutions at a protein surface

Andrew A. Pakula* & Robert T. Sauer

Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

* Present address: Division of Biology, California Institute of Technology, Pasadena, California 91125, USA

It is rare for amino-acid substitutions on the surface of proteins to have large stabilizing or destabilizing effects¹. Nevertheless, one substitution of this type, the Tyr 26 → Cys mutation in λ Cro, increases the melting temperature of the protein by 11 °C and the stability by 2.2 kcal mol⁻¹ (ref. 2). Here we show that the stability of Cro can be increased by many different amino-acid substitutions at position 26, with increasing stability showing a good correlation with decreasing side-chain hydrophobicity. As Tyr 26 is hyper-exposed to solvent in the Cro crystal structure^{3,4}, we suggest that wild-type and variant proteins with other hydrophobic side chains at position 26 are destabilized as a result of a reverse hydrophobic effect caused by the side chain being more exposed to solvent in the native than in the denatured state.

To investigate possible mechanisms by which amino-acid side-chain substitutions at position 26 alter the stability of λ Cro, we used cassette mutagenesis to generate six new substitutions at this site⁵. These variants, plus wild type and two previously isolated mutants, comprise a set of Cro proteins containing nine different residues (Asp, Cys, Gln, His, Leu, Phe, Trp, Tyr and Val) at residue 26. Each of these proteins behaves identically during purification, and the circular dichroism spectrum of each mutant is extremely similar to that of wild-type Cro, indicating that none of the substitutions causes a gross alteration of the native structure of Cro.

Representative melting curves for Cro proteins with substitutions at position 26 are shown in Fig. 1. Values of T_m and $\Delta\Delta G_u$ for each of the proteins are listed in Table 1. Polar residues such as Asp, His and Gln are among the most stabilizing substitutions, and aromatic residues such as Trp, Phe and Tyr are the least stabilizing. As shown in Fig. 2, increasing protein stability correlates reasonably well with decreasing side-chain hydrophobicity, as measured by transfer free energies from octanol to water^{6,7}.

Kauzmann⁸ proposed that the exposure of hydrophobic side chains to aqueous solvent results in an unfavourable free energy change by decreasing solvent entropy, presumably because water molecules form organized structures around non-polar side chains that are exposed to solvent⁹. During folding, many non-polar side chains become sequestered in the core of the folded

TABLE 1 Stability changes associated with residue substitutions at position 26 in λ Cro

Protein	T_m (°C)	$\Delta\Delta G_u$ (kcal mol ⁻¹)*
Trp 26	37.5	-0.1
Tyr 26	39.5	0.0 wild type
Phe 26	41.5	+0.4
Val 26	44.5	+0.9
Leu 26	46.0	+1.1
Gln 26	47.0	+1.4
His 26	49.5	+1.9
Cys 26	51.0	+2.2
Asp 26	54.0	+2.7

* $\Delta\Delta G_u$ values (± 0.2 kcal mol⁻¹) are calculated as $-RT \ln [K_u^{\text{mutant}}/K_u^{\text{wt}}]$, using values of K_u measured directly at 45 °C or calculated by extrapolation using the van't Hoff equation. As the unfolding reaction is a concerted transition from the folded dimer to two unfolded monomers², the $\Delta\Delta G_u$ values represent contributions from substitutions in each monomer. T_m values are ± 0.5 °C. Data for wild-type Cro and the Cys 26 mutant are from ref. 2.

protein, and the release of solvent from these side chains provides a large driving force for protein folding. But favourable hydrophobic effects are only operative for side chains, like those that form the protein core, that become less accessible to solvent in the folded protein. A non-polar side chain that is equally exposed to solvent in the native and denatured states should neither contribute to nor detract from stability. Moreover, some non-polar side chains may be less exposed to solvent in the denatured state than in the native state. For such residues, there would be a reverse hydrophobic effect which would oppose folding. Hence, if hydrophobic considerations are to account for our results, then residue 26 must be more solvent-accessible in native Cro than it is in denatured Cro.

Solvent accessibility of an amino-acid side chain in a protein is usually calculated relative to the accessibility of the same side chain in an Ala-X-Ala tripeptide in an extended reference conformation¹⁰. Hence, a side chain with a fractional solvent accessibility of 1.0 has the same exposure to solvent as the side chain in the reference tripeptide. At position 26 in Cro, the wild-type tyrosine side chain has a fractional solvent accessibility that is 1.4-fold greater than tyrosine in the reference tripeptide: this unusually high solvent exposure seems to result from its location at a turn in Cro, where the backbone conformation directs the Tyr 26 side chain away from main chain and neighbouring side-chain atoms^{3,4}. If the denatured state of Cro is a random coil, then the average fractional accessibility of a typical side chain should be close to 1.0. If denatured Cro is more compact than a true random coil, then the average side-chain accessibility should be slightly less than 1.0. In either model, it is likely that in the ensemble of denatured forms of Cro, the Tyr 26 side chain will experience an average solvent exposure

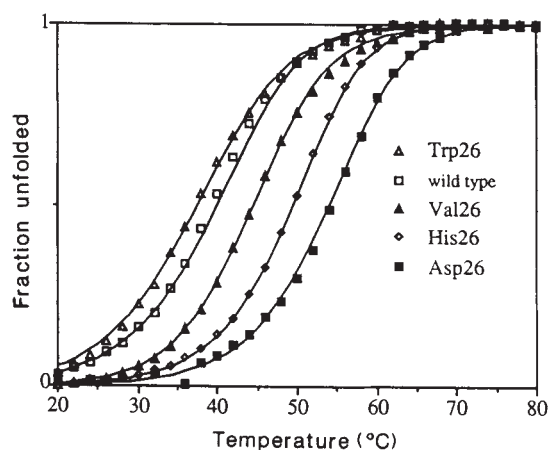


FIG. 1 Melting curves for wild-type Cro and variant Cro proteins with substitutions at positions 26. The continuous lines represent theoretical unfolding curves generated using the Gibbs-Helmholtz relation and best-fit values of ΔH and ΔC_p . Denaturation experiments used protein at concentrations of 120 $\mu\text{g ml}^{-1}$ in 10 mM potassium phosphate (pH 7.0), 0.1 M KCl, with temperature increased at a rate of 0.8 °C per min. Unfolding was monitored by circular dichroism ellipticity at 222 nm. The transition from the folded dimer (N_2) to two unfolded monomers (2U) is a concerted reaction in which the folded monomer is an unpopulated intermediate². The equilibrium constant for unfolding is defined as $K_u = [U]^2/[N_2]$.

METHODS. Mutations were introduced in the Cro gene of plasmid pAP119 (ref. 2). Mutant and wild-type proteins were purified from lysates of *Escherichia coli* (strain X9T, bearing overproducing plasmids¹⁶) by ion exchange chromatography on DEAE-Sephacel (Pharmacia) and Cellex P (Bio-Rad) and gel filtration chromatography on Sephadex G50 (Pharmacia) as described².

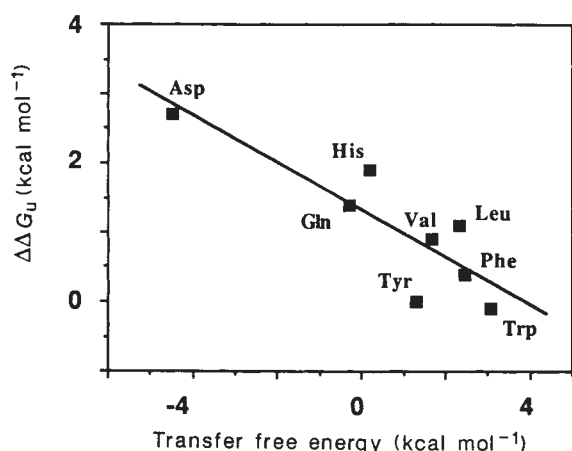


FIG. 2 Correlation between $\Delta\Delta G_u$ of unfolding for wild type and position 26 variants and side-chain hydrophobicity, represented by free energies of transfer from octanol to water^{6,7}. A point for the Cys26 protein is not included because a transfer free energy for cysteine is not available for this system. The transfer free energy used for Asp represents transfer of the uncharged Asp side chain from the organic to the aqueous phase ($\Delta G_t = -1.05 \text{ kcal mol}^{-1}$), followed by deprotonation of the side chain at pH 7.0 ($\Delta G = -3.4 \text{ kcal mol}^{-1}$). This value represents a minimum estimate for Asp hydrophobicity. The maximum estimate for Asp hydrophobicity would be represented by ΔG_t for the uncharged Asp side chain. The correlation between the stability of the position-26 Cro mutants and the free energy of transfer is not affected significantly by the choice of a hydrophobicity value for Asp. The correlation coefficient for the data shown is 0.74; the correlation coefficient obtained using the value for uncharged Asp is 0.70. This reflects the fact that Asp, in either the protonated or unprotonated form, is less hydrophobic than the other side-chain substitutions at position 26.

substantially lower than it would in folded Cro. Thus, a non-polar residue at this position could detract from stability through a reverse hydrophobic effect; that is, during protein folding the solvation of the Tyr 26 side chain would increase, giving rise to an unfavourable free energy term. If mutant substitutions at this position experience similar changes in solvation, then the more hydrophilic substitutions should be stabilizing because the reverse hydrophobic effect would be reduced.

Could mechanisms other than a reverse hydrophobic effect account for the observation that Cro is less stable when position 26 is hydrophobic than when it is hydrophilic? One possibility is that hydrophobic groups at position 26 decrease the stability of native Cro by stabilizing an aggregated form of unfolded Cro, but several observations indicate that this is not the case. First, the thermal denaturation of wild-type Cro is freely reversible², which would be surprising if unfolded Cro were aggregated and precipitated. Second, if unfolded Cro aggregates, then the folded protein should show reduced stability at higher protein concentrations. In fact, Cro stability increases with increasing protein concentration, which is to be expected for a transition from a folded dimer to two unfolded monomers². Finally, if hydrophobic aggregation of unfolded Cro caused the observed stability differences between the position-26 Cro variants, then these differences should be less pronounced in urea denaturation, where hydrophobic forces are reduced, than in thermal denaturation, where hydrophobic forces increase as the temperature increases. In urea denaturation experiments, however, the Tyr 26 and Asp 26 Cro proteins half denature at 1.8 M and 3.9 M urea, respectively, and the $\Delta\Delta G_u$ values vary from 3.1 to 3.4 kcal mol^{-1} over the range of 1.5 to 4.5 M urea (data not shown). These $\Delta\Delta G_u$ values are slightly larger than those measured for the same proteins by thermal denaturation (Table 1). Taken together, these considerations make it unlikely that

the effects of the position-26 substitutions are mediated by effects on the aggregation of denatured Cro.

The correlation between protein stability and side-chain hydrophobicity shown in Fig. 2 suggests that hydrophobicity is an important factor in determining the relative stabilities of Cro proteins substituted at position 26. Nevertheless, other factors such as hydrogen bonds and electrostatic interactions may also be operative for some substitutions. For example, the most stabilizing substitutions at position 26 are Asp and Cys, and the ionized forms of both side chains could interact favourably with the partial positive charge at the N-terminal end of the α -helix that begins at residue 27 of Cro^{3,4}. Such effects have been observed for acidic residues near the N termini of α -helices in T4 lysozyme¹¹.

Our results demonstrate that the presence of hydrophobic rather than hydrophilic groups on a protein surface can alter the equilibrium ratio of native to denatured protein to a significant extent. As discussed, reverse hydrophobic effects should only be operative for positions that are more exposed in the native than in the denatured protein. How common are such positions in other proteins? On the one hand, side chains that are as hyperexposed to solvent as Tyr 26 in Cro are extremely rare, and thus positions subject to reverse hydrophobic effects might also be expected to be rare. On the other hand, recent results indicate that some denatured proteins may be more compact than expected for a random coil^{12,13}. In these cases, reverse hydrophobic effects could arise for side chains having accessibilities near 1.0 in the native protein but lower in the denatured form. Large hydrophobic side chains appear to be excluded from many surface positions in the globin family¹⁴ and at several positions on the surface of λ repressor¹⁵. It is possible that reverse hydrophobic effects contribute to the restrictions at some of these positions. □

Received 19 July; accepted 19 December 1989.

- Hecht, M. H., Sturtevant, J. M. & Sauer, R. T. *Proc. natn. Acad. Sci. U.S.A.* **81**, 5685–5689 (1984).
- Pakula, A. A. & Sauer, R. T. *Proteins* **5**, 202–210 (1989).
- Anderson, W. F., Ohlendorf, D. H., Takeda, Y. & Matthews, B. W. *Nature* **290**, 754–758 (1981).
- Ohlendorf, D. H., Anderson, W. F., Fisher, R. G., Takeda, Y. & Matthews, B. W. *Nature* **298**, 718–723 (1982).
- Reidhaas-Olson, J. F. & Sauer, R. T. *Science* **241**, 53–57 (1988).
- Fauchere, J.-L. & Pliska, V. *Eur. J. Med. Chem. Chim. Ther.* **18**, 369–375 (1983).
- Eisenberg, D. & McLachlan, A. D. *Nature* **319**, 199–203 (1986).
- Kauzmann, W. *Adv. Prot. Chem.* **14**, 1–63 (1959).
- Tanford, C. *The Hydrophobic Effect* 2nd Edn (Wiley-Interscience, New York, 1980).

- Lee, B. K. & Richards, F. M. *J. molec. Biol.* **55**, 379–400 (1971).
- Nicholson, H., Becktel, W. J. & Matthews, B. W. *Nature* **336**, 651–656 (1988).
- Shortle, D. & Meeker, A. K. *Biochemistry* **28**, 936–944 (1989).
- Plitsyn, O. B. *J. Prot. Chem.* **6**, 273–293 (1987).
- Bashford, D., Choithia, C. & Lesk, A. M. *J. molec. Biol.* **196**, 199–216 (1987).
- Reidhaas-Olson, J. F. & Sauer, R. T. *Proteins* (in the press).
- Bowie, J. U. & Sauer, R. T. *J. biol. Chem.* **264**, 7596–7602 (1989).

ACKNOWLEDGEMENTS. We thank Brian Matthews for providing refined coordinates for λ Cro, and Jim Bowie, Kevin Shoemaker and Wendell Lim for comments and help with solvent accessibility calculations. This work was supported by the NIH.